Community Petascale Project for Accelerator Science and Simulation

A Proposal Submitted to the DOE Office of Science

Program Announcements:	DE-PS02-07ER07-09 and LAB 07-09
Program Area:	Accelerator Science and Simulation
Program Offices:	Offices of High Energy Physics, Nuclear Physics, and Basic Energy Sciences and the Computational Science Research and Partnerships (SciDAC) Division
Technical Contacts:	Glen Crawford, Craig Tull, Sidney A. Coon, Roger Klaffky, and Walter Polansky

Lead Institution: Fermi National Accelerator Laboratory.

Coordinating Principal Investigator and DOE Contact: Panagiotis Spentzouris
Fermi National Accelerator Laboratory
P.O. Box 500; Batavia, IL 60510-500
phone: 630-840-4342; FAX: 630-840-2783; spentz@fnal.gov

Participating Institutions and Principal Investigators

Argonne National Laboratory:	Michael Borland*, Lois Curfman McInnes*‡
Brookhaven National Laboratory:	Alexei Fedotov, Wolfram Fischer*
Fermi National Accelerator Laboratory:	James Amundson, Panagiotis Spentzouris*‡
Lawrence Berkeley National Laboratory:	William Fawley*, Esmond Ng*‡, Ji Qiang*, Robert Ryne*‡
Los Alamos National Laboratory:	Peter Walstrom*
Oak Ridge National Laboratory:	Richard Barrett*
Stanford Linear Accelerator Center:	Lie-Quan Lee, Cho Ng*‡
State University of New York, Stonybrook:	Jim Glimm*
Tech-X Corporation:	David L. Bruhwiler, John R. Cary*‡
Thomas Jefferson National Accelerator Facility:	Rui Li*, Haipeng Wang
University of California, Davis:	Kwan-Liu Ma*
University of California, Los Angeles:	Viktor Decyk, Warren Mori*‡, Sven Reiche*
University of Maryland:	Tom Antonsen*
University of Southern California:	Tom Katsouleas*

* Institutional Principal Investigator
‡ Management Committee

Community Petascale Project for Accelerator Science and Simulation

Project Summary

The DOE program of scientific discovery relies heavily on particle accelerators, which comprise 14 of the 28 facilities in the DOE twenty-year outlook on Facilities for the Future of Science. This proposed project, submitted to the Offices of High Energy Physics (HEP), Nuclear Physics (NP), Basic Energy Sciences (BES) and the Office of Advanced Scientific Computing Research (ASCR), will develop a comprehensive computational infrastructure for accelerator modeling and optimization. This project will advance accelerator computational capabilities from the terascale to the petascale to support DOE priorities for the next decade and beyond.

The SciDAC1 accelerator project, a partnership of accelerator computationalists, applied mathematicians, and computer scientists, generated a suite of parallel accelerator simulation tools. These were applied to important accelerator projects of the DOE. Under SciDAC2, these tools will be enhanced to contain new capabilities as needed by HEP projects, such as the ILC, the LHC, the Tevatron, and PEP-II, and for Advanced Acceleration research; NP projects, such as CEBAF and RHIC, the CEBAF and RHIC upgrades, RIA, and an NP electron collider, including ELIC and eRHIC; and BES projects, such as LCLS, NSLS-II, SNS, and upgrades to the APS.

This simulation suite will contain comprehensive set of interoperable components for beam dynamics, electromagnetics, electron cooling, and advanced accelerator modeling. Beam dynamics studies will include developing an understanding of the lifetime limits from beam collisions in colliders. Electromagnetics modeling will be used to optimize cavity shapes for increased accelerating gradient and beam current. Electron cooling computations will determine the configuration of cooling systems needed for mitigating beam-beam effect. Advanced accelerator modeling is needed to develop concepts for HEP accelerators beyond the ILC and to develop tabletop electron accelerators for BES and NP projects.

In each of these areas, the modeling tools require petascale supercomputers and advanced software for making effective use of these large, parallel platforms. Computational infrastructure in the areas of shape determination and optimization; advanced adaptive meshing; dynamic load balancing; embedded boundaries; component methodologies; performance measurement, assessment and improvement; linear and nonlinear solvers; and visualization will be used and advanced. Consequently, critical to this effort will be the embedded collaborations with the applied mathematics and computer science communities. In the end, high-quality computational tools developed with the best computational physics, applied mathematics, and computer science will be made available to the US particle accelerator community through installation at government laboratories, universities, and industry.

Contents

Title Page i

Project Summary i

Table of Contents ii

Project Narrative 1

1 Executive Summary and Motivation 1

2 Past Accomplishments 3
- 2.1 Codes Developed under the AST Project . . . 3
- 2.2 Application Highlights . . . 4

3 Proposed Research 7
- 3.1 Development of Modeling Capabilites . . . 7
- 3.2 Target Applications of the Program Offices . . . 11

4 Budget and Project Milestone Overview 13

5 Management Plan 14

References 16

A Committees and Collaboration list 22

B Petascale 24

C Target Applications of the Program Offices 26

D X-Ray Light Source Modeling for Ultrafast Science 30

E Embedded Scientific Application Partnerships 34
- E.1 Development of Novel Scalable Parallel Eigensolvers for Large Electromagnetic Cavity Simulations . . . 35
- E.2 Shape Determination and Optimization for Superconducting Cavities . . . 38
- E.3 Domain Specific Scalable Linear Solvers for Finite-Element Large-Scale Accelerator Cavity Simulations . . . 41
- E.4 Development of a Super-Fast Poisson Solver for Accelerator Physics Applications . . . 44
- E.5 Parallel Adaptive Refinement for High-Fidelity Finite-Element Electromagnetic System Simulations . . . 47
- E.6 Embedded Boundary Methods for Particle-in-Cell Simulations of Electromagnetic Structures . . . 50
- E.7 The UPIC Framework . . . 53
- E.8 Performance Analysis and Optimization of Accelerator Modeling Codes on Petascale Supercomputers . . . 56
- E.9 Interactive and Remote Visualization for Massive Complex Field-Particle Data on Unstructured Grids . . . 59

E.10 High-Performance Parallel Data Management and Analysis Tools for Beam Dynamics Modeling . . . 63
E.11 Computational Quality-of-Service Enabled Beam Dynamics Modeling . . . 67

F Summary of Tasks and Milestones for Each Institution 71
F.1 Work Plan – Fermi National Accelerator Laboratory . . . 72
F.1.1 Tasks Supporting the Office of High-Energy Physics . . . 72
F.1.2 Tasks Supporting the Office of Advanced Scientific Computing Research . . . 74
F.2 Work Plan – Argonne National Laboratory . . . 75
F.2.1 Tasks Supporting the Office of High-Energy Physics . . . 76
F.2.2 Tasks Supporting the Office of Nuclear Physics . . . 76
F.2.3 Tasks Supporting the Office of Basic Energy Sciences . . . 77
F.2.4 Tasks Supporting the Office of Advanced Scientific Computing Research . . . 78
F.2.5 Details on Argonne Tasks Supporting the Office of Nuclear Physics . . . 79
F.3 Work Plan – Lawrence Berkeley National Laboratory . . . 81
F.3.1 Tasks Supporting the Office of High-Energy Physics . . . 81
F.3.2 Tasks Supporting the Office of Nuclear Physics . . . 85
F.3.3 Tasks Supporting the Office of Basic Energy Sciences . . . 86
F.3.4 Tasks Supporting the Office of Advanced Scientific Computing Research . . . 88
F.4 Work Plan – Los Alamos National Laboratory . . . 91
F.4.1 Tasks Supporting the Office of High-Energy Physics . . . 91
F.5 Work Plan – Oak Ridge National Laboratory . . . 94
F.5.1 Tasks Supporting the Office of Advanced Scientific Computing Research . . . 94
F.6 Work Plan – Stanford Linear Accelerator Center . . . 95
F.6.1 Tasks Supporting the Office of High-Energy Physics . . . 96
F.6.2 Tasks Supporting the Office of Nuclear Physics . . . 97
F.6.3 Tasks Supporting the Office of Basic Energy Sciences . . . 98
F.6.4 Tasks Supporting the Office of Advanced Scientific Computing Research . . . 99
F.7 Work Plan – Tech-X Corporation . . . 101
F.7.1 Tasks Supporting the Office of High Energy Physics . . . 101
F.7.2 Tasks Supporting the Office of Nuclear Physics . . . 102
F.7.3 Tasks Supporting the Office of Advanced Scientific Computing Research . . . 103
F.8 Work Plan – University of California at Davis . . . 104
F.8.1 Tasks Supporting the Office of Advanced Scientific Computing Research . . . 104
F.9 Work Plan – UMD . . . 105
F.9.1 Tasks Supporting the Office of High Energy Physics . . . 105
F.10 Work plan – BNL . . . 106
F.10.1 Tasks Supporting the Office of Nuclear Physics . . . 106
F.11 Work Plan – Stony Brook University . . . 108
F.11.1 Tasks Supporting the Office of Advanced Scientific Computing Research . . . 108

G Biographical Sketches 109

H Institutional Facilities and Resources 200
H.1 Argonne National Laboratory Facilities . . . 200
H.2 Brookhaven National Laboratory Facilities . . . 202
H.3 Fermi National Accelerator Laboratory Facilities . . . 203
H.4 Lawrence Berkeley National Laboratory Facilities . . . 204

H.5 Oak Ridge National Laboratory Facilities 205
H.6 Stanford Linear Accelerator Center Facilities 206
H.7 Stony Brook University Facilities 207
H.8 Thomas Jefferson National Accelerator Facility Facilities 208
H.9 Tech-X Corporation Facilities 209
H.10 University of California at Davis Facilities 210
H.11 UCLA Facilities 211

I Collaboration Agreement Letters 212

1 Executive Summary and Motivation

Particle accelerators are critical to scientific discovery in the DOE program in America and indeed the world [1]. Of the 28 facilities listed in the DOE report, "Facilities for the Future of Science: A Twenty-Year Outlook," 14 involve accelerators. The development and optimization of accelerators is essential for advancing our understanding of the fundamental properties of matter, energy, space, and time, and for enabling research in materials sciences, chemistry, geosciences, and aspects of biosciences. In the 15-year plan [2] for High Energy Physics (HEP), the first two action items call for full support of the program of the LHC at CERN, and for establishing leadership in the R&D effort to design and build the proposed ILC on U.S. soil. At the same time, it is imperative to maximize the physics reach of the ongoing DOE/HEP program, and that involves the performance optimization of current accelerators such as the Fermilab Tevatron and the SLAC B-Factory.

The Nuclear Physics (NP) program utilizes accelerators to study the properties of nuclear matter, the structure of the nucleus, understand the mechanism of quark confinement, and create and study a heretofore-unknown state of nuclear matter, the quark-gluon plasma. The flagship NP accelerators are CEBAF at Jefferson Lab and RHIC at BNL. The worldwide nuclear physics community has identified [3] the construction of a polarized electron-ion collider as a long-term objective in the on-going effort to answer these questions. DOE/NP is considering two alternative approaches for such a facility: e-RHIC [4] and ELIC [5]. DOE/NP is also considering a proposed rare isotope facility that will permit studies of nuclei far from stability that promise to improve radically our understanding of atomic nuclei.

The flagship accelerator facilities for the Basic Energy Sciences (BES) program are its spallation neutron sources (SNS, LANSCE) and its synchrotron light sources (ALS, APS, NSLS, SSRL). Over the past decade, the light-source community consensus is that short wavelength free-electron lasers (FEL's) will form the basis for the so-called fourth generation light sources. DOE/BES has committed to building the LCLS at SLAC (at a cost of ≈400M$) aiming to enhance our understanding of the arrangement of atoms in inorganic and organic materials and in biological molecules, thus enabling the development of new materials and molecules with desirable properties.

SciDAC1 produced a powerful suite of parallel simulation tools representing a paradigm shift in computational accelerator science. Simulations that used to take weeks or more now take hours, and simulations once thought impossible are now performed routinely. These codes were applied to important DOE projects including the Tevatron, PEP-II, LHC, RHIC, NLC, ILC design, SNS, and LCLS [6, 7, 8, 9]. They were also applied to advanced concepts such as plasma-based accelerators, where they played a key role in understanding the physics of doubling the energy of a 42 GeV beam at SLAC and of low-energy-spread beam production in laser wakefield accelerators. The former result will be published in Nature [10], which devoted a feature issue to the latter [11, 12], as well as listing it as one of the top 10 discoveries of 2004.

Under SciDAC2—recognizing the complexity, precision, and beam intensity requirements of next generation accelerators—our paradigm will change from single machine, single-component simulations to end-to-end (multi-stage or complete system), multi-physics simulations. Building upon the foundation laid under SciDAC1, we will extend our terascale capabilities to the petascale, and add new capabilities to deliver a comprehensive, fully integrated accelerator simulation environment. Our team adds new institutions (ANL, ORNL, Stony Brook, TJNAF) to the previous SciDAC1 collaboration. Our proposed modeling applications involve nearly every accelerator laboratory in the country. Collaboration with ASCR researchers will help ensure that this effort will utilize the most appropriate algorithms for future petascale platforms. Furthermore, it will help provide the infrastructure to accomplish a key objective of the present proposal: the *integration* of the high-end capabilities that we will develop.

Throughout the project we will maintain close contact with our target applications, both to aid in the application of our simulation tools and also to ensure that we are aligned with the priorities of DOE/SC and its scientific requirements. Our management structure will establish a formal mechanism to ensure that our project remains on track.

The applications of the new capabilities we will develop will tackle the most computationally challenging problems of near, medium, and long-term priorities of HEP, NP, and BES. This includes large-scale electromagnetic modeling of a three-cryomodule RF unit of the ILC, with realistic cavity shapes and misalignments; assessment of the impact of wakefields on beam dynamics; and multi-physics, multi-bunch modeling of ILC beam dynamics, especially in the damping rings, to assess beam loss and emittance growth, to study their effect on luminosity, and for design optimization and cost reduction. We will also focus on design optimization of accelerator components with complicated geometries, including the hybrid RIA RFQ and ILC crab cavity which includes couplers with very fine features. For the LHC we will focus on beam-beam and electron-cloud simulations to help understand and optimize machine performance. For the Tevatron we will focus on beam-beam and impedance effects to help understand the anti-proton current limitations. For the RHIC II proposal, the e-RHIC concept design, the CEBAF upgrade proposal and the ELIC concept design, we will focus on three areas: 1) electromagnetic simulation of superconducting RF cavities, with and without self-consistent beam treatment, 2) multi-physics beam dynamics simulations with emphasis on nonlinearities, beam-beam effects, and intrabeam scattering, and 3) electron cooling physics, aiming to quantitatively understand the dynamical friction force on ions moving through electron distributions in the presence of strong external fields. In the area of accelerator-based X-ray light sources, our petascale tools will be applied to: 1) understanding and predicting limits to beam brightness, coherent and incoherent undulator radiation, emittance preservation, and microbunching 2) start-to-end simulations for the design of future facilities, and 3) coupling our codes to control systems to support commissioning of near-term facilities like LCLS.

Our project will emphasize the interaction of beam dynamics and electromagnetics codes and applications. For example, in the ILC the wakefields obtained by electromagnetic modeling will be the input for beam dynamics end-to-end simulations to determine beam quality. The results will be fed back to cavity builders for design improvement. This iterative procedure involving accelerator builders and designers and incorporating beam dynamics and electromagnetic simulations at the petascale will ensure successful accelerator design.

In addition, the project will assist the development of advanced accelerator concepts. These technologies have already demonstrated gradients and focusing forces more than 1000 times greater than conventional technology. SciDAC1 codes were used to predict and explain these results. Under SciDAC2, we will provide real-time or near-real-time feedback between simulation and advanced accelerator experiments. We will answer several grand physics questions of importance to the HEP and BES offices. For HEP, we will determine whether $\frac{2}{3}$ of an ILC beam can be used to drive a plasma wave wake that doubles the energy of the other $\frac{1}{3}$ in meter distances while maintaining beam quality and luminosity, and we will determine whether short-pulse, high-power lasers or particle beams can be used to make compact 1 GeV to 1 TeV electron/positron accelerators. For BES we will determine if plasma accelerators can produce ultra-short (< 10 fs), ultra-bright electron beams for a coherent light source beyond LCLS. Applying petascale resources to these concepts may open a new path to the ultra-high energy frontier and, at the same time, lead to a new generation of compact accelerators for advancing US science, technology, industry, and medicine.

In summary, the Community Petascale Project for Accelerator Science and Simulation (COMPASS), a comprehensive and coordinated effort involving HEP, NP, and BES, will advance and integrate capabilities in beam dynamics, electromagnetics, and advanced accelerator concept modeling, for accelerator design, analysis, and discovery.

2 Past Accomplishments

The present proposal builds upon the successes of the SciDAC1 Accelerator Science and Technology (AST) project, both by extending our present terascale capabilities to the petascale and by developing new capabilities to meet the needs of HEP, NP, and BES programs. Here we summarize past accomplishments of the AST project in regard to code development and code applications.

2.1 Codes Developed under the AST Project

The AST project involved 3 main thrust areas for accelerator modeling: Beam Dynamics (BD), Advanced Accelerators (AA), and Electromagnetics (EM). The parallel 3D simulation codes developed under SciDAC1, or with partial support from SciDAC1, for these areas include [8, 13, 14, 15, 16]:
(1) A comprehensive suite of parallel 3D PIC codes for beam dynamics modeling

BeamBeam3D - A code for simulating weak-strong and strong-strong beam-beam interactions with a variety of collision geometries [17]. It uses a hybrid decomposition for optimal parallel performance. The code has been successfully benchmarked against VEPP2 data [18].

IMPACT - A code suite widely used for modeling high intensity beams in RF proton and electron linacs and photoinjectors. Selected features include Poisson solvers for a variety of boundary conditions, energy binning, high aspect ratio solvers, wakefield effects, and a 1D CSR model.

MaryLie/IMPACT (ML/I) - An extensible, multi-physics framework that combines the nonlinear optics capabilities of MaryLie 5.0 with the 3D space-charge and acceleration capabilities of IMPACT. The code allows for map production, map analysis, particle tracking, 3D envelope tracking, fitting, and optimization, all within a single, coherent user environment [16].

Synergia - A multi-language, extensible framework utilizing state-of-the-art numerical libraries, solvers, and physics models. Synergia features 3D space-charge and impedance modules, and arbitrary order Lie maps for magnetic optics. Selected features include multi-bunch, ramping and RF and magnet, multi-turn injection, and active feedback modeling. It utilizes multiple Poisson solvers including an FFT-based solver from IMPACT and a multigrid solver [19].
(2) A versatile 3D parallel PIC code suite for plasma acceleration, e-cloud, and e-cooling

UPIC - A Framework which includes components for electrostatic, electromagnetic and Darwin field solvers in multi-dimensions. It has dynamic load balancing for fields and particles and higher order splines for accurate energy conservation. It supports mixed shared and distributed memory architectures [20]. It can be used to develop efficient e-cooling code.

OSIRIS - A fully electromagnetic code that includes higher order splines, current and field smoothing for accurate energy conservation, Vay PML open boundary conditions, a relativistically correct two-body collision model, and ionization. It has been validated against experiment.

QuickPIC - The only 3D parallel PIC code for modeling plasma accelerators in the quasi-static limit, also capable of electron cloud modeling in conventional accelerators [21]. QuickPIC was built using UPIC modules. It has been validated against experiment and benchmarked with OSIRIS.
(3) A comprehensive suite of parallel 2D/3D electromagnetic modeling codes using higher-order (up to 6th order) finite elements on unstructured grids for geometry fidelity and field accuracy [22]:

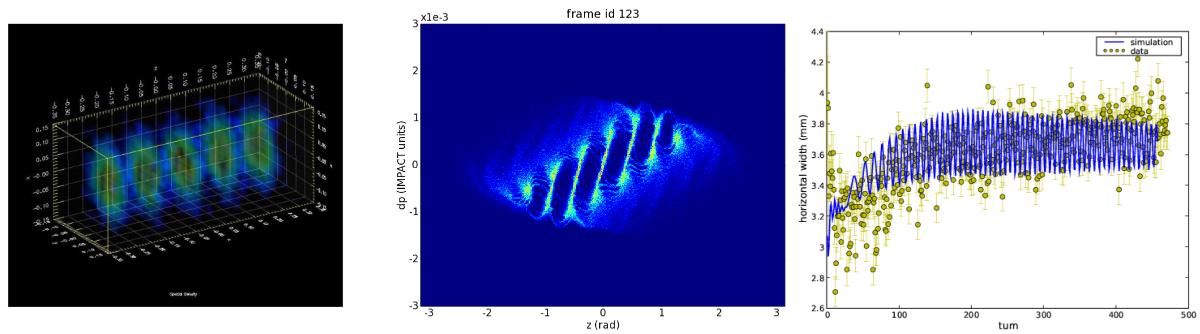
Omega3P - A 3D eigensolver for treating lossless, lossy, periodic and waveguide-loaded cavities.
S3P - A 3D frequency domain solver to calculate S parameters of RF components.
T3P - A 3D time-domain solver with beam for wakefield computations.
Track3P - A 3D particle tracking code with surface physics for simulating dark current and multipacting.
Pic2P - A 2D particle-in-cell code for space-charge dominated device simulations.

The suite includes V3D, a visualization package for viewing particles and fields in 3D complex geometries.

(4) VORPAL: A parallel framework for modeling electromagnetic fields, fluids, particles, and their interactions, using a structured mesh approach, which allows for conformal, embedded boundaries for modeling curved surfaces and boundaries [23, 15]. The VORPAL framework exploits multiple modeling techniques, including particle in cell, explicit and implicit electromagnetics, electrostatics, particle-in-cell, fluids, surface processes, such as field-induced and secondary emission and sputtering, and molecular dynamics. Effective modeling of electron cooling is possible [24] through use of the molecular dynamics methods with operator splitting, a semi-analytic model for binary electron-ion collisions, and electrostatic PIC [25] to incorporate long-range electron-electron interactions and arbitrary external fields.

2.2 Application Highlights

Beam dynamics



(a) Merging of 5 linac microbunches in the FNAL booster.

(b) Longitudinal phase space shows halo and space-charge "drag" during bunch merge.

(c) Beam profile evolution compared with Ionization Profile Monitor experimental data.

Figure 1: Fermilab Booster Synergia simulation, including 3D space-charge, 2nd order maps, multi-turn injection, and rf ramping models.

The widespread usage of our SciDAC1 codes in the accelerator community is a testament to our success in developing and distributing a new generation of beam dynamics modeling tools. Their application has led to a number of successes, detailed in [7], and to a number of important "firsts" in computational accelerator simulation:

- First million particle, million turn, strong-strong colliding beam simulation for LHC [26]
- First multi-bunch, multi-turn injection simulation from linac-to-booster w/ self-consistent 3D space charge [19]
- First 100M simulation of a linac for an x-ray light source w/ self-consistent 3D space charge [27]

AST accomplishments in the beam dynamics area are documented in more than 25 refereed publications and numerous conference proceedings. Selected examples of successful applications of our codes include:

LHC - We modeled electron-cloud emittance dilution effects using QuickPIC in long-range (1000's of turns) simulations. Also, we established that the single-kick approximation commonly used before our studies produces erroneous results [28]. This application is an excellent example of the synergistic development promoted by our project: QuickPIC was originally developed to enable fast simulations of plasma-based accelerators, but under SciDAC1 it was enhanced to be applicable to circular accelerator problems.

Fermilab Booster - Using Synergia, we performed the first-ever simulation of the process of linac microbunch capture, debunching, and acceleration, including beam position feedback models, all using a 3D space-charge model, (Figures 1(a) and 1(b)). The results of our simulations were compared to beam measurements, both under normal machine operations and during machine studies (Figure 1(c)). These simulations provided guidance to operators to reduce losses and maximize Booster intensity and commission the Booster collimators [29, 30].

ILC Damping Ring - The MaryLie/Impact framework was used to model space-charge effects in the ILC Damping Ring (DR). Our team worked closely with the ILC researchers and made specific enhancements to ML/I for these studies. ML/I was one of the main codes used in the studies that led to the selection of the damping ring design detailed in the Baseline Configuration Document of the ILC.

Tevatron BeamBeam3d was used for the first strong-strong beam-beam simulations of the Tevatron including a lattice model based on measurements, and an impedance model.

Future Light Sources & High Brightness Electron Beams - Though the AST project was originally funded by the HENP program, AST codes have proven to be very useful to BES. For example, Figure 2 shows results from an IMPACT simulation of the microbunching instability in an x-ray light source using 2M to 100M particles and multiple physical phenomena [27]. More generally, high brightness electron beams are important to many projects across DOE/SC. Under SciDAC1, IMPACT was applied to several photoinjector projects nationwide including photoinjectors at ANL, BNL, Cornell, FNAL/NIU, TJNAF, LBNL, and SLAC/LCLS.

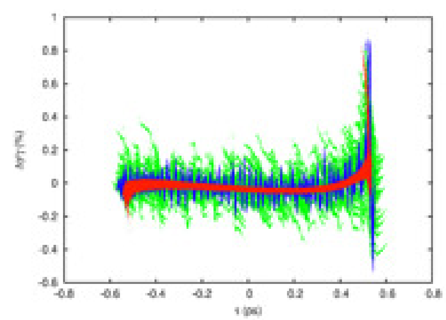


Figure 2: Microbunching instability in a linac-driven x-ray light source. 2M particles (green) and 100M particles (blue) with space-charge and no-CSR, compared with no-space-charge, no-CSR (red).

Electromagnetics

The SciDAC1 accomplishments in the EM area of the AST project have been documented in several SciDAC reports [6, 31, 23]. Major accomplishments are highlighted below.

PEP-II - Beam heating and its power deposition profile in the PEP-II interaction region (IR) were computed. Redesign for the IR upgrade led to higher beam current, higher luminosity and particle discovery. *TSTT* partitioning schemes sped up the simulations by a factor of 10.

Next Linear Collider (NLC) - Work included cell design, first ever end-to-end wakefield simulation of a prototype 55-cell structure, and discovery of pulse rise time effects on dark current. Using Omega3P, the cell was modeled to the required frequency accuracy of 0.01%, validating this more efficient design that would have saved over $100M in machine's cost [32]. *TOPS* provided a direct solver that improved the solution time by 2 orders of magnitude.

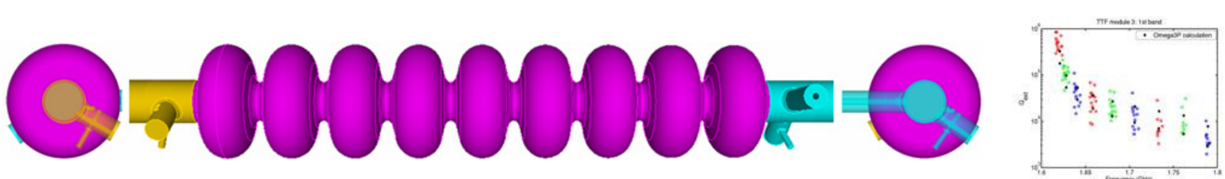


Figure 3: The 9-cell TTF cavity and the Qe comparison between measurements (deformed shape) and Omega3P calculations (ideal shape) of the 1st dipole band.

ILC - First-ever Omega3P wakefield computations of the entire Tesla cavity (Fig. 3) and identification of trapped modes in a cryomodule (Fig. 4), successful prediction of multipacting barriers

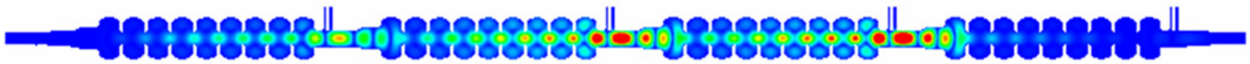


Figure 4: Omega3P calculation of a trapped HOM in the ILC ICHIRO 4-cavity superstructure.

observed in the Low-Loss cavity using Track3P [33], detailed computations of power flow at HOM couplers using T3P/VORPAL (Fig. 5) for corroborating with beam position measurements [34], and VORPAL detailed beam map calculations for tracking in crab cavity.

PERC helped optimize Omega3P performance on Cray X1E at ORNL's NCCS [35]. Collaborations with UC Davis on V3D visualization led to the discovery of mode rotation in the Tesla cavity. With *TOPS* and *TSTT*, a new automatic capability to replace manual, iterative process for optimizing cavity design has started.

CEBAF - First successful simulations of the TJNAF niobium cylindrical TE011 mode test cavity using VORPAL showed that the experimentally observed frequency split between TE011 and TM111 modes is due to the symmetry breaking of the input couplers. Also, using Omega3P, the dipole modes responsible for beam breakup in the 7-cell superconducting cavity were identified.

RHIC - First simulation of an RF photocathode gun including detailed treatment of the emission process was carried out. This represents the only implementation of the detailed photocathode model of Jensen and coworkers [36] in a self-consistent 3D electromagnetics code.

RIA - The frequencies of RFQ cavities were computed to higher accuracy than standard codes which will allow for reduced number of tuners and ease operating procedures [37, 38]. *TSTT* tools on adaptive mesh refinement significantly improved the convergence to required accuracies.

LCLS - The LCLS RF gun cavity was designed entirely through simulation to meet beam dynamics requirements and to address pulse heating effects [39]. A fabricated prototype performed well as designed. Lately, efficient and accurate PIC simulations of the RF gun showed the effects of wakefield and retardation in this space-charge dominated device [14].

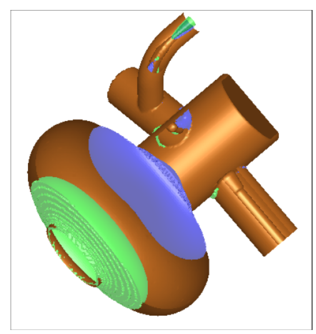


Figure 5: Power flow out an HOM damper.

Electron cooling

RHIC-II - Use of 4th-order predictor-corrector algorithm [40, 41] in VORPAL resolved uncertainties in and between asymptotic and parametric models for the friction force in magnetized electron coolers. This resolution influenced the conceptual design for the cooler of the RHIC-II luminosity upgrade. Development and use of a new semi-analytic algorithm for Coulomb collisions in VORPAL [24] showed that wiggler-magnet induced electron motion, intended to decrease recombination, would only weakly reduce the friction force. This work was critical in the decision to change the conceptual design of the RHIC-II cooler from solenoid-based to wiggler-based.

Advanced accelerators

The accomplishments of the advanced accelerator effort under SciDAC1 have led to the publication of ten Physical Review Letters [42, 43, 44, 45, 46, 47, 48, 49, 50, 51], being highlighted in the Fall 2006 volume of SciDAC Review [52], in the publication of four articles, including one cover picture, in Nature [53, 54, 55, 10]. The main author of QuickPIC, Chengkun Huang, received the 2006 best thesis prize from the APS Division of Computational Physics and Cameron Geddes received the best thesis prize from the Division of Plasma Physics for his LBNL experiments modeled with VORPAL.

Successful applications include modeling the major US plasma-based accelerator experiments at SLAC and LBL, developing key physics understanding of PWFA and LWFA in the nonlinear blowout regime, and applying plasma-based codes to e-cloud interactions in conventional accelerators. In some cases, these codes predicted important effects in plasma acceleration, later to be verified by experiment, and helped guide the design of the experiments.

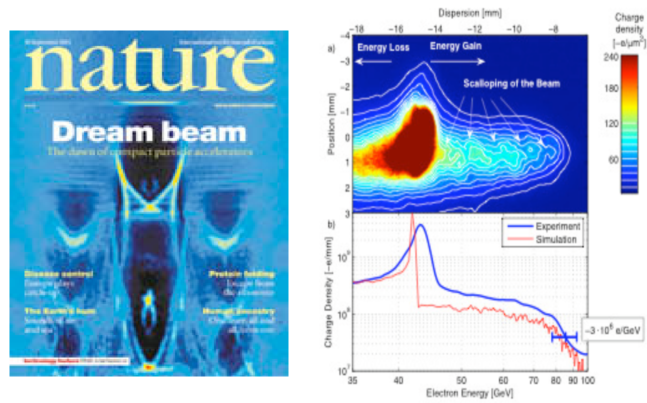


Figure 6: SciDAC codes were used to model LWFA and PWFA experiments that have been or will be published in Nature. Left: Image on cover from a VORPAL simulation of a LWFA experiment. OSIRIS simulation was used in another article in this issue. Right: Measured and simulated (using QuickPIC) energy spectrum from a PWFA experiment at SLAC in which the energy of 42 GeV electrons was doubled.

3 Proposed Research

Our vision is to develop a comprehensive suite of interoperable accelerator modeling tools which will enable *Virtual Prototyping* (VP) of the design and optimization of accelerator components, and multi-physics *Virtual Accelerator Modeling* (VAM), for use by accelerator scientists across DOE/SC. This goal can only be achieved by developing tools and applications capable of utilizing petascale computing. A detailed discussion on the need for petascale capabilities appears in appendix B.

For each of the DOE offices HEP, NP, and BES, we have identified high priority modeling problems by gathering input from representatives of the program offices and their accelerator projects.

3.1 Development of Modeling Capabilites

Beam dynamics

In order to validate designs of today's complex accelerators, multi-physics modeling is essential. We will design, implement, test, and optimize beam dynamics software modules for multiple physical phenomena, and incorporate them into our accelerator modeling framework.
(1) *Space-charge effects* - We will port our SciDAC1 solvers, including Integrated Green Function solvers [9, 56], to petascale platforms, and we will develop new solvers (see Appendix E.4).
(2) *Beam-beam effects* - Our SciDAC1 codes will be extended to include: full nonlinear symplectic dynamics, a rotating beam model with a crossing angle, quantum effects, a wire compensation model, and ability to model linac-ring colliders.
(3) *Multi-species effects: Electron-cloud, Fast-Ion* - We will use two complementary approaches to e-cloud and fast-ion modeling, one based on WARP/POSINST [57] and another based on QuickPIC [21], and interface e-cloud generation modules with e-cloud dynamics modules. We will improve both the physics models and performance, potentially increasing scalability by up to 100×.
(4) *Intrabeam scattering* - We will implement (i) a fast, stochastic map-based approach with effective dynamical friction coefficients and diffusion tensor calculated from non-self consistent models, and (ii) a self-consistent approach using the Rosenbluth potential with a grid-averaged distribution [58]. These models will be benchmarked with data from RHIC and other models [59].
(5) *Beam-environment interactions: beam loading and wakefield effects* - Our approach involves models of varying approximation: (a) extend our SciDAC1 model (precalculated wakefields) to include single bunch longitudinal and transverse wakes in a single cavity, (b) use precalculated

eigenmodes in a time-dependent equation that includes external and beam-induced cavity excitation, and (c) utilize VORPAL for self-consistent particles-fields calculation.
(6) *High order optics, dynamic errors and feedback* - Mature high order optics models are already utilized in our codes. We will maintain these and implement improved models of soft-edge magnets, including wigglers, and a new kicker cavity model. We will also extend the multi-bunch capabilities of our frameworks to enable full modeling of multi-bunch instabilities and include models of dynamically changing quantities (jitter, position), errors, and feedback systems.
(7) *High brightness e-beam and radiation production physics* - Current approaches to CSR modeling involve 1D approximations or time-consuming 3D models that would be a bottleneck in large scale ($\geq$ 100M particle) simulations with parallel PIC codes. We will develop a parallel 3D CSR solver using a Mesh-Green function approach along with parallel particle and field managers for performance and scalability. We will also explore new multi-scale approaches for improved fidelity in simulations of high-brightness beams for X-ray FELs (XFELs) and other applications. Since start-to-end design of XFELs also requires modeling of X-ray production, we will establish interfaces between our beam dynamics codes and the FEL design codes GINGER, GENESIS, and SPUR. We will improve the parallelization of XFEL codes with a view toward scaling to $\geq 10^4$ processors, and develop advanced field algorithms for XFEL simulation.

Electromagnetics

The design and optimization of accelerator components for both present and future accelerators requires detailed electromagnetic simulation capability, which over the past decade, has matured greatly due to diverse applications. Some of the issues that we can address with our electromagnetic codes are:
(1) *Design and optimization of cavities* - Superconducting RF (SRF) cavities for ILC, CEBAF, RIA and SNS are designed and optimized to satisfy requirements on field gradient, HOM damping and cryogenic loss. For RFQs in RIA, constraints on the accelerating frequency and Ohmic wall loss require simulations of high fidelity and accuracy in order to save operating cost. To improve the accuracy and efficiency, we will develop adaptive h-p-q refinement [60, 61, 62] in the finite element method and higher-order embedded algorithms in the structured grid method, respectively. Furthermore, we will implement a multi-mode waveguide boundary condition [63] in the finite element eigensolver for handling multiple waveguide mode propagation.
(2) *Wakefields and HOMs* - Trapped modes in cavities will not only cause multi-bunch effects on beam emittances, but potential heating effects as well. We will develop solvers that can handle trapped modes in structures at the system level. Short-range wakefields generated by beamline components may cause single bunch instabilities, and the resulting HOM power generated will propagate and heat up vulnerable components such as bellows in the vacuum chamber. For accurate determination of short-range wakefields for short bunches, we will implement and compare a moving beam window using a non-dispersive algorithm [64] in VORPAL and higher-order finite elements in T3P. We will also implement a Napoly-type wakefield integration technique [65, 66] for collimators or intrusive structures in the vacuum chamber.
(3) *Multipacting and dark current* - Multipacting during high power processing can prevent a cavity from reaching its designed gradient, e.g., the ILC ICHIRO cavity [33]. Dark current from surface field emissions will alter the RF properties of a cavity and generate unwanted backgrounds downstream in particle colliders. We will develop efficient algorithms for handling particle emission and location at higher-order surfaces to add to the already well-tested tracking algorithms.
(4) *Beam driven device modeling* - Self-consistent PIC simulations are required for devices dominated by space charge, including RHIC SRF guns, the ILC L-band sheet beam klystron and most

ERLs. We will develop a 3D finite element PIC code, Pic3P, based on the 2D PIC code Pic2P [14]. The charge-conserving property of the embedded boundary algorithm [67] in VORPAL will enhance the PIC handling capabilities at curved surface. We will validate the two codes with each other for cases including regimes that presently do not have known solutions.
(5) *Integrated EM/thermal/mechanical modeling* - Thermal and mechanical considerations are as important as electromagnetics in cavity design. The RF power loss on the walls will be used as input to the thermal solver to find the temperature gradient, which will then be input to the mechanical solver to calculate the wall distortions. The development of this integrated tool will facilitate the design of RF windows and the analysis of Lorentz force detuning in SRF cavities [68].

Electron cooling

We will model the ion velocity drag (*i.e.*, friction force) in electron coolers for high-energy colliders like RHIC-II, eRHIC and ELIC, to verify the coolers' efficacy and help optimize the design. Due to the extreme difficulty of such calculations and the lack of experimental data for high-energy systems, we will establish confidence in these calculations by pursuing three complementary approaches that will be used to cross-check one another: molecular dynamics, PIC and a Langevin approach.

We will dramatically expand our VORPAL molecular dynamics simulations [40, 41, 24] from a few ions in a small, periodic domain up to the full transverse extent of the e^- and ion beams, retaining periodicity only in the longitudinal direction. The planned algorithm developments will increase the effective overlap of computation with communication, speed-up the treatment of ion-ion collisions, more effectively suppress the diffusive ion dynamics as compared to velocity drag, and follow the spirit of PPPM [69] approaches for much faster treatment of distant e^-/ion collisions.

We will explore use of a highly optimized PIC code based on UPIC with very fine grid resolution and with a Particle-Particle-Particle-Mesh (PPPM) [69]. Aliasing errors will be reduced by use of quadratic interpolation and filtering. The module in PPPM for direct binary interactions can be replaced by reduced methods to model particle collisions, such as described in Takizua [70].

We will also use a Langevin approach [71] to solve the Fokker-Planck equation self-consistently using the Rosenbluth potential (or Landau integral) [72, 73]. We pioneered this approach under SciDAC1 and demonstrated its feasibility [58]. We will extend the technique to electrons and ions, include effects of magnetic fields [74, 75, 76, 77], explore an efficient FMM-based technique to compute F-P collisional operator, implement improved time integration schemes for electrons and ions, and integrate the F-P code with our other modules for multi-physics modeling.

Advanced accelerators

The vision is to develop the capability of modeling future 100 GeV to 1 TeV stages with proper resolution and to model existing 1 GeV to 50 GeV experiments with real-time feedback. As described in Appendix B, we need petascale simulations to achieve this vision. Below we describe how we intend to ensure that the parallel PIC tools scale to and run efficiently on such machines.

High energy PWFA stages cannot currently be modeled with standard PIC codes. The key difficulty for modeling a 0.5 TeV PWFA afterburner stage is that the radius of the beam (sub-μm) is small compared to the accelerating structure ($\sim$50μm), requiring thousands of grid points in each transverse direction. QuickPIC is structured with a 2D slice of plasma moving through a drive beam at each 3D time step. Thousands of longitudinal grid points are required. The 2D solve at each longitudinal slice currently scales to $\sim$1024 processors. In order for the code to scale beyond this a software pipelining algorithm [78]needs to be implemented which exploits the fact that the front of the drive beam can be advanced in time as soon as the 2D plasma slice passes by it. In principle, N separate copies of the code operating simultaneously after the N^{th} 3D time step has been reached where based on preliminary tests we expect this to allow the code to 100,000+ processors. Greater scaling will be achieved if the field solve can be more efficient.

The challenge for modeling higher energy stages using full PIC codes is that the number of time steps needed increases faster than the memory needed. Therefore, the codes will need strong and not weak parallel scaling where only $\sim$1000 particles might exist on each processor. Therefore, the development of optimal particle managers for the different hardware platforms and precise load balancing algorithms will be essential for the scalability PIC codes. We will also implement, optimize and use higher-order field and particle updates, new fluid algorithms, and ponderomotive guiding center (PGC) PIC [79] algorithms primarily in VORPAL.

Another goal is to provide the tools for real time steering of SLAC and LOASIS experiments by the end of the project. In order to realize this goal, the turn-around time needs to be reduced to less than one hour. Currently a QuickPIC simulation of a 30 GeV electron or positron beam propagating through 1 meter of a gas takes $\sim$2000 hours. A VORPAL or OSIRIS simulation of a 1 GeV LWFA stage in the blowout regime takes $\sim$70,000 node hours, a full PIC run of future LOASIS experiments (5cm long plasmas) will take $\sim 2,500,000$ node hours, and preliminary QuickPIC simulation of the LOASIS experiments take $\sim$2500–5000 node hours. We assume that 100,000 to 1,000,000 processor computers will be available during the lifetime our project. The goal of real-time steering can be met using QuickPIC by a combination of increasing the single processor speed by a factor of 5-10, reducing the number of iteration loops to $\sim$2, scaling the non pipelined version from 32 to 100+ processors, and adding pipelining for another factor of $\sim$256.

Applied Mathematics and Computer Science Support

Mathematical techniques, advanced scalable numerical algorithms, and computational tools are important ingredients in COMPASS. While many of the mathematical and computational tools we will employ are relatively mature, we must advance their capabilities and develop new approaches to meet the petascale computational challenges facing the project. Building on the collaborations and successes in SciDAC1, our team will work closely with many SciDAC CETs and Institutes (listed in Appendix E) to deploy state-of-the-art developments and to advance application-specific modeling capabilities in accelerator physics. This section briefly highlights some of the key issues and approaches that we are exploring under the umbrella of embedded Scientific Application Partnerships (SAPS); further details are presented in Appendices E and I.

Solvers, Meshing, Optimization, and Particle Methods. As discussed in Section 3, trapped-mode analysis of waveguide-loaded SRF cavities in an RF unit consisting of three cryomodules is a computationally intensive task in the ILC Computational Grand Challenge, requiring the solution of large-scale nonlinear eigenvalue problems. Working closely with TOPS [80], we will develop novel and scalable algorithms for solving these eigensystems; see Appendix E.1 for details.

One of the ultimate goals of the ILC Computational Grand Challenge is to automatically perform shape determination and optimization, which is posed as a numerical optimization problem with PDE constraints. As discussed in Appendix E.2, the solution requires a well-coordinated effort between domain scientists and researchers from various areas in ITAPS [81] and TOPS.

Extremely large sparse linear systems arise in shape determination and shape optimization, as well as in driven frequency computations, time-domain simulations, and eigenvalue calculations. As discussed in Appendix E.3, we will work closely with TOPS to explore novel, robust, and scalable direct and iterative linear solvers.

Fast and scalable Poisson solvers have important applications in many accelerator beam dynamics simulation codes, as described in Section 3. FFT-based methods are widely used but do not scale well because of global communication. As discussed in Appendix E.4, we will investigate a new super-fast Poisson solver, which is based on direct solvers for sparse linear systems but exploits the numerical low rank structures inherent in the Poisson operator.

Complicated cavity geometries pose another issue in the simulation of electromagnetic systems. Parallel adaptive mesh refinement in element size (h), order of basis functions (p), and order of geometry surface (q) is essential. We will work closely with ITAPS and CSCAPES [82] to develop adaptive (h, p, q) refinement strategies, including high-order error estimators and load balancing techniques; more discussion is in Appendix E.5.

A complementary approach to electromagnetics for complex structures is to use embedded boundary methods with regular meshes, where elements of the field update stencil are modified in cells cut by the boundary. To achieve end-to-end accelerator modeling, there is a need for higher-order embedded boundary algorithms that can be combined with interior algorithms with better dispersive properties, and for more accurate interpolation within cut cells for incorporating particle effects. As discussed in Appendix E.6, we will work with ITAPS to meet these goals.

An essential part of COMPASS is the extension of UPIC's [20] PIC capabilities to support scalable advanced accelerator and beam dynamics modeling via QuickPIC [21]. In collaboration with APDEC [83], we will develop a faster and more robust quasi-static solver, devise optimized dynamic load balancing strategies for 3D beam computations, and introduce capabilities in APDEC's Chombo for structured AMR and field solvers. As discussed in Appendix E.7, this work will improve the efficiency and scalability of UPIC as well as QuickPIC and MaryLie/IMPACT [84].

Performance Analysis. A common theme of COMPASS accelerator codes is the requirement for effective petascale computation, which requires exploiting complex details at all levels of architectural design. As discussed in Section E.8, we will work with PERI [85] to apply advanced optimization techniques that take full advantage of petascale architectures in order to achieve significant runtime reductions, with initial emphasis on BeamBeam3D, IMPACT, Omega3P, and T3P.

Visualization and Data Analysis. Advanced visualization and data analysis are vital parts of the design and optimization of accelerators, including advancing the understanding of complex processes that arise in their operation in new parameter regimes. Thus, as discussed in Appendices E.9 and E.10, we will collaborate with the IUSV Institute [86], VACET CET [87], and SDM CET [88] to introduce new visualization and data management techniques to facilitate scientific discovery through images and animation of massive amounts of field and particle data.

Integration. Software infrastructure for multi-physics accelerator modeling is a central part of COMPASS. Our approach builds upon ongoing work on Synergia2 and MaryLie/IMPACT and will extend to other subdomains of the project. Using the Common Component Architecture [89, 90], we will devise common interfaces that encapsulate preexisting physics modules, thereby facilitating code reuse, the incorporation of new capabilities, and the exploration of performance tradeoffs among various algorithms and implementations for different simulation scenarios and target petascale machines. We also will collaborate with TASCS, PERI, and TOPS to provide support for computational quality of service to help select, parameterize, and reconfigure COMPASS codes dynamically during runtime; further details are in Appendix E.11.

3.2 Target Applications of the Program Offices

In this section we summarize the COMPASS application activities. Appendix C contains a detailed discussion of these activities.

HEP applications: We will provide integrated beam dynamics modeling capabilities in support of LHC commissioning, with emphasis on fully nonlinear optics, beam-beam and electron-cloud modeling.

For the ILC Global Design Effort we will model the wakefield effects for a single RF unit consisting of three cryomodules, taking into account realistic cavity imperfections and misalignments (100s of μm). We will incorporate our newly enhanced capabilities in a new finite-element based code TEM3P, an integrated thermal/electromagnetic/mechanical multi-physics solver, to form a complete engineering prototyping toolset. The wakefields from these simulations will be used in our beam dynamics codes to study multi-bunch effects on emittance dilution. For the damping rings, we will evaluate the short-range wakefields and provide a broadband impedance budget as input for beam dynamics instability studies, identify all the trapped modes and optimize performance and cost for all the vacuum chamber components. We will also perform multi-bunch, multi-physics beam dynamics studies, including electron cloud, impedance, IBS, higher-order magnetic optics, CSR, and space-charge in the same model. Ultimately, we aim to perform simulations of the full cycle of the machine, from injection to extraction. For alternative cavity designs, such as the Low-Loss and the ICHIRO cavities, we will focus on optimization of the end-groups, the notch gap fields, and the coax pickup power flow. We will also model multipacting in the ICHIRO cavity, to help maximize gradient. Wakefields from cavity designs will be input into beam dynamic codes to study x-y coupling due to 3D effects such as cavity deformations and mode rotation. For the Beam Delivery System design we will focus on models of wakefield effects from collimators, crab cavities and the interaction region (IR). We will also study emittance dilution and beam instabilities in the low emittance transport, the main linac, and the IR, using the multi-physics, multi-bunch model described above, with the addition of full beam-beam modeling, including quantum effects.

For the Tevatron we will develop 3D strong-strong beam-beam simulations of the motion of colliding bunches in the Tevatron, including impedance and optics parameters reflecting the measured lattice and helical orbit. We will study possible coherent instabilities arising from the combined action of the beam-beam and impedance effects.

We will perform e-cloud generation and dynamics simulations for parameters relevant to the Main Injector (MI) proton plan, and compare the predictions with MI measurements. In addition, we will perform 3D beam dynamics simulations of the recycler, including multi-particle effects, such as space-charge, for the recycler parameters envisioned by the super-beams for NuMI plan (SNuMI).

We will use our mature model of the Fermilab Booster to implement a prototype Control Room application. We will develop an interface to machine parameter and detector readouts to implement an almost-real-time controls algorithm to help maximize intensity and efficiency. The application can them be ported to other machines of the FNAL proton source.

We will support the DOE advanced accelerator facilities such as SABER and LOASIS advanced accelerator facilities by providing codes and simulation support for designing experiments, for full-scale modeling of ongoing experiments, and eventually for real time steering. In addition, we will perform simulations to determine the feasibility of 100 to 500 GeV stages.

NP applications: We will work on optimizing the normal-conducting Radio Frequency Quadrupole (RFQ) and the superconducting TEM class resonators [91, 92, 93] the Rare Isotope Accelerator (RIA) Facility. In the injector section of the RIA, post-accelerator focusing by RF fields can save millions of dollars compared with alternative methods. For the optimization, integrated electromagnetic and mechanical analysis is required due to the complexity of the cavity geometry and mechanical properties.

For the CEBAF 12 GeV upgrade we will utilize our integrated RF/thermal/stress capability to design high-power RF windows for low temperature structure cryo-cooling. Simulations of multipacting and dark currents will help identify problems during the high power processing of the superstructures, and those of space charge dominated effects will help evaluate the beam quality at the end of the injector.

The design challenge of the BNL SRF-photoinjector is to produce a 5 to 10 nC charge, high brightness bunch with good emittance. To support this effort, we will perform 3D PIC simulations including wakefields, image charges, and retarded potentials with a large number of electrons to resolve short bunches (10–20 ps) starting off from the cathode non-relativistically.

For the RHIC upgrade, eRHIC, and ELIC we will perform electron cooling simulations in order to optimize design parameters. Since the parameter space is large, we will use task-farming techniques to facilitate efficient parameter scans on emerging architectures with 10^5 processors. We will also simulate the electron cooling ERLs to study the effects on machine efficiency from energy recovery of heated electrons, by utilizing our self-consistent EM capabilities.

In order to help improve the performance of the current RHIC operation, guide the design efforts of the planned RHIC upgrades, and enhance confidence on the proposed eRHIC and ELIC design, we will perform simulations of colliding beams, utilizing models of a linac-ring and a circulator ring-ring collider, a crab cavity, a wire compensator, an electron lens, IBS, and the space-charge effect of parallel co-moving beam. The new capabilities will be tested against analytical results and experimental data from RHIC.

BES applications: We will use simulations including peak field and multipacting in the HOM coupler of the SNS SRF cavities, to understand feed-through failures in the end-groups. In addition, we will model the effects of HOM damping to determine if the HOM couplers are really necessary for the cavities envisioned for the SNS power upgrade.

We will perform feasibility studies for the APS upgrade [94] Energy Recovery Linac (ERL) option, to answer questions regarding beam quality and gun sustained current [95], and wakefield effect minimization through 3D profile optimization to maximize current. Also, many of the electromagnetic capability development for the ILC, including HOM buildup and feedback onto the beam, could be applied to the APS upgrade.

We will extend our simulation capability for start-to-end modeling of accelerator-based X-ray sources for ultrafast science. We will develop advanced parallel capabilities and new models for simulating phenomena such as CSR and X-ray production, and use them to explore limits to beam brightness, emittance preservation, and to study the microbunching instability. An important component of this work is to couple the models to light source control rooms. As an early application, we will give LCLS control room scientists a smoothly operational suite of start-to-end (i.e., cathode to FEL user station) computational tools able to model the entire system based on actual lattice settings, giving near real-time predictions of diagnostic output and FEL performance.

4 Budget and Project Milestone Overview

Our proposed support is \$4.465M for accelerator physics goals of the HEP, NP, and BES with ASCR support for applied mathematics and computer science. This involves 14 institutions: 8 national labs, 5 universities, and one private research company and ensures the education of new computational scientists and technology transfer. It achieves synergism through strong connections to other projects and efficiency through integration. Nevertheless, different offices have different primary targets, thus permitting this breakdown.

We propose \$2.035M for HEP developments and applications. Roughly 35% of the effort is for electromagnetic (EM) modeling, with nearly all of this work for ILC applications, including SRF acceleration cavities, the beam delivery system, damping rings, and klystrons. The main participants in this work are SLAC and Tech-X with contributions from FNAL. Roughly 46% of the HEP effort is for beam dynamics (BD), with the main participants being LBNL and FNAL, and with effort at SLAC, UCLA and USC. The priority targets of HEP/BD are to support LHC commissioning and the ILC Global Design Effort, followed by Tevatron Run II operations, Main Injector proton plan, and SNuMI. All seven of the BD code development areas listed in Sec. 3.1

impact ILC, while the other applications involve up to six of these. Roughly 19% of the HEP effort will be in advanced acceleration (AA), with primary participation from UCLA and Tech-X and secondary participation from LBNL, USC, and the U. of Maryland.

The project proposes \$1.05M in support of developments and applications of NP. Roughly 43% of this will be for EM, including cavities and SRF guns. The primary efforts will be by SLAC and Tech-X with some effort at TJNAF, BNL, and ANL. Roughly 33% will be for electron cooling (EC) computations, in support of RHIC-II, ELIC, and eRHIC. Code development will be carried out at Tech-X, LBNL, and UCLA. Finally, 26% of effort will support advanced modeling of colliding beams at in RHIC, eRHIC, and ELIC. Primary efforts will be at LBNL, with effort also at BNL and TJNAF.

The project proposes \$290K in support of developments and applications with specific relevance to BES. The majority (\$215K) will support a startup effort involving ANL, LBNL, SLAC, and UCLA on advanced modeling of accelerator-based X-ray sources for ultrafast science. The remainder (\$75K) will support EM modeling at SLAC in support of SNS and APS-II. The X-ray effort will grow in year 2 and year 4 as it as it progresses from startup to a mature effort.

Support of \$1.09M is proposed for applied math and computer science specific to this project. This will ensure efficient application of ASCR funded software to accelerator computation. This includes solvers, meshing, adaptive refinement, algorithm development, performance optimization, component technologies, quality of service, and visualization. Institutions participating in this part of the project are SLAC, LBNL, ANL, UCLA, FNAL, Stony Brook, Tech-X, ORNL, and UC Davis, in descending order of effort.

In Appendix F we describe the tasks, scheduling, milestones and FTE assignments for each institution. Our project involves both near term milestones, to provide time critical modeling capabilities, and long term development tasks. The application delieverables are listed in section 3.2. The long term goal of this project is to produce interoperable components and infrastructure for a comprehensive, start-to-end, parallel framework for accelerator design. We will pursue this task over the full 5 years of the project, together with the continous optimization and adaptation of our algorithms and solvers to the petascale.

5 Management Plan

The strategy of the COMPASS project is to develop a comprehensive set of interoperable accelerator physics modeling components capable of utilizing petascale computing and to serve the needs of the highest priority applications of HEP, NP, and BES. Our organizational approach provides strong management focus on, and accountability for, the execution of our plans. The COMPASS management structure emphasizes coordination with the SciDAC CETs and Institutes, to ensure utilization of the most advanced computational tools in our software development, and with representatives of our target applications to ensure optimal utilization of our modeling tools.

Overall responsibility for this effort will be vested in the COMPASS Management Committee, whose members (Cary, Mori, McInnes, C. Ng, E. Ng, Ryne, Spentzouris) are also Principal Investigators. The coordinating-PI (Spentzouris) serves as the Chair of the committee. The Management Committee sets the project's goals and milestones, and it devises the plans for meeting them. It sets priorities, decides on fund allocation, and ensures that work is completed on schedule. The committee is responsible for coordinating with the ASCR centers and the target applications. At the end of each project year, it develops a rolling two year road map for both the modeling capabilities development and the target applications. The institutional work packages with milestones and initial FTE assignments are summarized in Appendix F. Dr. Bakul Banerjee of FNAL will prepare and maintain the WBS of the project. Schedule slips for project milestones will be reported to the Management Committee, and the committee will decide whether reallocation of resources or a change of scope is necessary.

The Management Committee holds at least one conference call per month and communicates via email between calls. Decisions are made by consensus; when consensus is not reached, decisions are made by majority vote, with the chair's vote deciding the outcome in case of a tie. The chair of the Management Committee (Spentzouris) serves as the principal contact with the DOE. Each institution receiving funds under this grant has a Principal Investigator. The institutional representatives assume responsibility for the work carried out at their institution and submits quarterly reports to the Management Committee. The coordinating-PI will submit quarterly progress reports to DOE and is responsible for tracking the overall grant budget. He will be assisted in preparing these reports and in tracking the grant budget by Bakul Bannerjee. The Management Committee will organize quarterly full collaboration meetings prior to the submission of quarterly reports.

As the SciDAC solicitation calls for open source software, COMPASS will include a number of mechanisms to disseminate output. Project software will be made publically available through a centralized web repository, which will also include a database of all available documentation. Installation assistance on supported supercomputers will also be made available. Extensive examples and validation tests to verify the correctness of installations will be distributed with each product. The project will organize several tutorials to facilitate community use of the products, possibly including at the U.S. Particle Accelerator School.

The project management structure and the breakdown of the level 2 activities is shown in Figure 7. The high-level project activities are divided into three categories: accelerator simulation software development (models and integration), SAP activities and coordination with the SciDAC CETs/Institutes, and software applications.

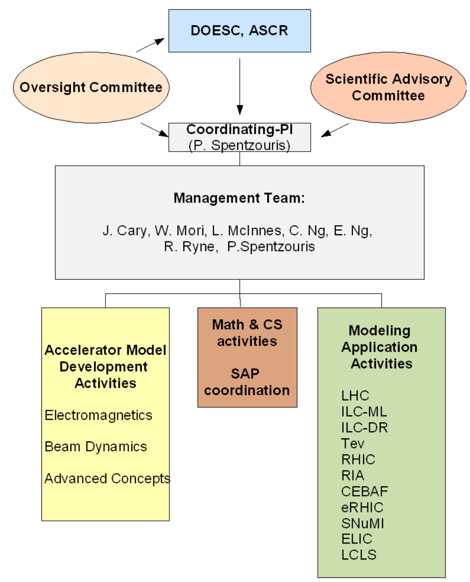


Figure 7: COMPASS management structure.

The COMPASS management plan includes an Oversight Committee (OC) and a Scientific Advisory Committee (SAC), which will assist the Management Committee in maintaining an up-to-date list of HEP, NP, and BES priorities and in managing project activities.

The SAC monitors the progress of the scientific applications of the project and provides leadership in setting priorities. The SAC consists of representatives of our stakeholders (the HEP, NP, and BES researchers) and experts in applied mathematics and computer science. The names of the committee members are listed in Appendix A. The SAC organizes an annual meeting, possibly inviting other experts if the need arises, to review progress and obtain input on future directions. The SAC also provides recommendations to the Management Committee to ensure that project resources are allocated to best serve the needs of the DOE/SC programs. The objective of this process is to achieve the greatest scientific benefit from the project's resources through broad input from the community. In addition to the annual meeting, the coordinating-PI can call for additional SAC meetings if additional input is needed.

The OC, with representatives from the management of the major lab participants (ANL, FNAL, LBNL, SLAC) and from the university community, is charged with monitoring the progress of the project's activities and assisting the DOE/SC in programmatic issues and scheduling reviews. The OC is also responsible for advising the Management Committee on coordination issues within our management plan.

References

[1] S. Abraham, Spencer Abraham announces department of energy 20-year science facility plan, text available at http://www.er.doe.gov/Sub/Newsroom/News_Releases/DOE-SC/2003/Energy%2020-Year%20Science20Facility%20Plan.htm, 2003.

[2] N. R. C. Committee on Elementary Particle Physics in the 21st Century, Revealing the hidden nature of space and time:charting the course for elementary particle physics, The National Academies Press, ISBN: 0309101948, 2006.

[3] A. Deshpande, R. Milner, R. Venugopalan, and W. Vogelsang, Ann. Rev. of Nuclear and Particle Science **55**, 0506148 (2000).

[4] e-RHIC website, http://www.bnl.gov/cad/eRHIC/.

[5] ELIC website, http://casa.jlab.org/research/elic/elic.shtml.

[6] K. Ko et al., Impact of SciDAC on accelerator projects across the Office of Science through electromagnetic modeling, in Proc. of SciDAC Conference, San Francisco, California, 2005.

[7] R. Ryne, J. Phys.: Conf. Ser. **46**, 171 (2006).

[8] P. Spentzouris, Accelerator modeling under SciDAC: meeting the challenges of next-generation accelerator design, analysis, and optimization, http://bel.gsi.de/icap2006/PAPERS/THMPMP03.PDF, 2006, 9th Intern. Comp. Accel. Physics Conf. (Invited Talk).

[9] J. Qiang, S. Lidia, R. Ryne, and C. Limborg-Deprey, Phys. Rev. ST/AB **9**, 044204 (2006).

[10] I. Blumenfeld et al., Energy Doubling of 42 GeV Electrons in a Metre Scale Plasma Wakefield Accelerator, Accepted by Nature.

[11] C. G. R. Geddess et al., Nature **431**, 538 (2004).

[12] S. Mangles et al., Nature **431**, 535 (2004).

[13] D. L. Bruhwiler, Parallel simulation of coulomb collisions for high-energy electron cooling systems, http://bel.gsi.de/icap2006/TALKS/WEA1MP01_TALK.PDF, 2006, 9th Intern. Comp. Accel. Physics Conf. (Invited Talk).

[14] A. Candel, Parallel higher-order finite element method for accurate field computations in wakefield and PIC simulations, http://bel.gsi.de/icap2006/PAPERS/WEMPMP03.PDF, 9th Intern. Comp. Accel. Physics Conf. (Invited Talk).

[15] J. R. Cary, High-performance self-consistent electromagnetic modeling of beams, http://bel.gsi.de/icap2006/TALKS/MOAPMP02_TALK.PDF, 2006, 9th Intern. Comp. Accel. Physics Conf. (Invited Talk).

[16] R. Ryne, Recent progress on the MaryLie/IMPACT beam dynamics code, http://bel.gsi.de/icap2006/PAPERS/TUAPMP03.PDF, 2006, 9th Intern. Comp. Accel. Physics Conf. (Invited Talk).

[17] J. Qiang, M. Furman, and R. Ryne, Phys. Rev. ST Accel. Beams **5** (2002).

[18] J. Amundson et al., J. Physics: Conf. Series **46**, 205 (2006).

[19] J. Amundson, P. Spentzouris, J. Qiang, and R. Ryne, J. Comp. Phys. **211**, 229 (2006).

[20] V. K. Decyk and C. D. Norton, Computer Physics Communications **164**, 80 (2004).

[21] C. Huang et al., Journal of Computational Physics **217**, 658 (2006).

[22] C. Ng et al., State of the art EM field computation, in Proc. of European Particle Accelerator Conference, Edinburgh, Scotland, Jun 26-29, http://accelconf.web.cern.ch/Accelconf/e06/PAPERS/THXFI01.PDF., 2006.

[23] J. R. Cary et al., J. Physics: Conf. Series **46**, 200 (2006).

[24] G. I. Bell et al., Simulations of dynamical friction including spatially-varying magnetic fields, volume 821, page 319, 2006.

[25] P. Messmer and D. L. Bruhwiler, Comp. Phys. Comm. **164**, 118 (2004).

[26] J. Qiang, Strong-strong beam-beam simulation on parallel computer, in ICFA Beam Dynamics Newsletter No. 34, 2004.

[27] I. Pogorelov, Recent developments in impact and application for future light sources, `http://bel.gsi.de/icap2006/PAPERS/WEPPP01.PDF`, 2006, 9th Intern. Comp. Accel. Physics Conf.

[28] G. Rumolo et al., Phys. Rev. ST Accel. Beams **6** (2003).

[29] P. Spentzouris and J. Amundson, Emittance dilution and halo creation during the first milliseconds after injection at the fermilab booster, in *AIP Conf.Proc.773*, pages 127–131, 2005.

[30] J. Amundson, W. Pellico, P. Spentzouris, T. Sullivan, and L. Spentzouris, Nucl. Instrum. Meth. **A570**, 1 (2007).

[31] L.-Q. Lee et al., Achievements in ISICS/SAPP collaborations for electromagnetic modeling of accelerators, in Proc. of SciDAC Conference, San Francisco, California, 2005.

[32] Z. Li et al., Parallel computations of wakefields in linear colliders and storage rings, in Proc. of Particle Accelerator Conference, Chicago, Illinois, June 18-22, 2001.

[33] A. Kabel, Accelerating cavity design for the International Linear Collider, `http://bel.gsi.de/icap2006/PAPERS/WEA4IS01.PDF`, 9th Intern. Comp. Accel. Physics Conf. (Invited Talk).

[34] S. Molloy and *et al.*, Phys. Rev. ST/AB **9**, 112802 (2006).

[35] K. J. Roche, Application software case studies in FY05 for the Mathematical, Information and Computational Sciences Office of the US Department of Energy, 2005.

[36] K. L. Jensen, D. W. Feldman, and P. G. O'Shea, App. Phys. Lett. **85**, 5448 (2004).

[37] L. Ge et al., Advanced eigensolver for electromagnetic modeling of accelerator cavities, in Proc. of the 15th Conference on the Computation of Electromagnetic Fields, Shenyang, China, June 26-30, 2005.

[38] L. Ge et al., Adaptive mesh refinement for high accuracy wall loss determination in accelerating cavity design, in Proc. of the 11th Biennial IEEE Conference on Electromagnetic Feild Computation, Seoul, Korea, June 6-9, 2004.

[39] L. Xiao et al., Dual feed rf gun design for the LCLS, in Proc. of Particle Accelerator Conference, Knoxville, tennessee, May 16-20, SLAC-PUB-11213, 2005.

[40] A. V. Fedotov et al., Phys. Rev. ST Accel. Beams **9**, 074401 (2006).

[41] A. Fedotov, I. Ben-Zvi, D. Bruhwiler, V. Litvinenko, and A. Sidorin, New J. Phys. **8**, 283 (2006).

[42] B. E. Blue et al., Physical Review Letters **90**, 214801 (2003).

[43] M. J. Hogan et al., Phys. Plasmas **7**, 2241 (2000).

[44] M. J. Hogan et al., Physical Review Letters **90**, 205002 (2003).

[45] M. Hogan et al., Physical Review Letters **95**, 054802 (2005).

[46] P. Muggli et al., Physical Review Letters **93**, 014802 (2004).

[47] D. K. Johnson et al., Phys. Rev. Lett. **97**, 175003 (2006).

[48] S. P. D. Mangles et al., Physical Review Letters **94**, 245001 (2005).

[49] S. P. D. Mangles et al., Physical Review Letters **96**, 215001 (2006).

[50] F. S. Tsung et al., Physical Review Letters **93**, 185002 (2004).

[51] W. Lu, C. Huang, M. Zhou, W. Mori, and T. Katsouleas, Phys. Rev. Lett. **96**, 165002 (2006).

[52] W. B. Mori, SciDAC Review , 16 (2006).

[53] S. P. D. Mangles et al., Nature **431**, 535 (2004).

[54] J. Faure et al., Nature **431**, 541 (2004).

[55] C. G. R. Geddes et al., Nature **431**, 538 (2004).

[56] J. Qiang, M. Furman, and R. Ryne, J. Comp. Phys. **198**, 278 (2004).

[57] J.-L. Vay, A. Friedman, and D. P. Grote, Self-consistent simulations of high-intensity beams and e-clouds with warp-posinst, `http://bel.gsi.de/icap2006/PAPERS/WEA3MP02.PDF`, 2006, 9th Intern. Comp. Accel. Physics Conf. (Invited Talk).

[58] J. Qiang, R. Ryne, and S. Habib, Self-consistent langevin simulation of coulomb collisions in charged-particle beams, in *Proceedings of SC2000*, 2000.

[59] V. A. Lebedev, Nucl. Instrum. Meth. **A391**, 176 (1997).

[60] S. Chen and L.-Q. Lee, Gradient-based higher order error indicator in accelerator cavity simulation, Presented at 2006 Bay Area Scientific Computing Day, 2006.

[61] D. Pardo, L. Demkowicz, C. Torres-Verdín, and L. Tabarovsky, International Journal for Numerical Methods in Engineering **65**, 1269 (2006).

[62] D. Pardo, L. Demkowicz, C. Torres-Verdín, and L. Tabarovsky, Computer Methods in Applied Mechanics and Engineering (In Press).

[63] L.-Q. Lee, Nonlinear eigenproblems in modeling International Linear Collider cavities, Presented at Fudan University, China, available at `http://www.slac.stanford.edu/~liequan/FuDan06_llee.pdf`, 2006.

[64] I. Zagorodnov and T. Weiland, Phys. Rev. ST/AB **8**, 042001 (2005).

[65] I. Zagorodnov, Phys. Rev. ST/AB **9**, 102002 (2006).

[66] H. Henke and W. Bruns, Calculation of wake potentials in general 3D structures, in Proc. of EPAC 2006, p2170-2172, 2006.

[67] R. Barker, N. Luhmann, J. Booske, and G. Nusinovich, *Modern Microwave and Millimeter Wave Power Electronics*, IEEE Press, 2005.

[68] G. Devanz, M. Luong, and A. Mosnier, Numerical simulation of dynamic Lorentz detuning of SC cavities, in Proc. of EPAC 2002, p2220-2222, 2002.

[69] R. W. Hockney and J. W. Eastwood, *Computer Simulation Using Particles*, Adam Hilger, 1988.

[70] T. Takizua and H. Abe, J. Comp. Phys. **25**, 205 (1977).

[71] H. Risken, *The Fokker-Planck Equation: Methods of Solution and Applications*, Springer, 1989.

[72] L. D. Landau, Phys. Z. Sowjetuion **10**, 154 (1936).

[73] M. Rosenbluth, W. M. MacDonald, and D. L. Judd, Phys. Rev. **107**, 1 (1957).

[74] N. Rostoker, Phys. Fluids **3**, 922 (1960).

[75] G. J. D. Montgomery, L. Turner, Phys. Fluids **17**, 954 (1974).

[76] T. M. O'Neil, Phys. Fluids **26**, 2128 (1983).

[77] K. Matsuda, Phys. Fluids **26**, 1508 (1983).

[78] C. Huang et al., J. Phys.: Conf. Ser. **46**, 190 (2006).

[79] P. Messmer and D. L. Bruhwiler, Phys. Rev. ST/AB **9**, 031302 (2006).

[80] David Keyes (PI), Towards Optimal Petascale Simulations (TOPS), SciDAC2 CET, `http://www.scidac.gov/math/TOPS.html`, 2006.

[81] Lori Freitag Diachin (PI), Interoperable Technologies for Advanced Petascale Simulations (ITAPS), SciDAC2 CET, `http://www.scidac.gov/math/ITAPS.html`, 2006.

[82] Alex Pothen (PI), Combinatorial Scientific Computing and Petascale Simulations (CSCAPES), SciDAC2 Institute, `http://www.scidac.gov/math/CSCAPES.html`, 2006.

[83] Phil Colella (PI), Applied Partial Differential Equations Center (APDEC), SciDAC2 CET, `http://www.scidac.gov/math/APDEC.html`, 2006.

[84] R. Ryne and et al., MARYLIE/IMPACT: A parallel beam dynamics code with space charge, http://scidac.nersc.gov/accelerator/marylielayer.html.

[85] Bob Lucas (PI), Performance Evaluation Research Institute (PERI), SciDAC2 Institute, http://www.scidac.gov/compsci/PERI.html, 2006.

[86] Kwan-Liu Ma (PI), SciDAC Institute for Ultrascale Visualization (IUSV), SciDAC2 Institute, http://www.scidac.gov/viz/ultraviz.html, 2006.

[87] E. Wes Bethel (PI), Visualization and Analytics for Enabling Technology (VACET), SciDAC2 CET, http://www.scidac.gov/viz/VACET.html, 2006.

[88] Arie Shoshani (PI), Scientific Data Management (SDM), SciDAC2 CET, http://www.scidac.gov/compsci/SDM.html, 2006.

[89] CCA Forum homepage, http://www.cca-forum.org/, 2006.

[90] D. E. Bernholdt et al., Intl. J. High-Perf. Computing Appl. **20**, 163 (2006), ACTS Collection special issue.

[91] P. N. Ostroumov, V. N. Aseev, and A. A. Kolomiets, Application of a new procedure for design of 325 MHz RFQ, submitted to J. Nucl. Instrum. Meth., 2006.

[92] P. N. Ostroumov, B. Mustapha, and J. A. Nolen, Computational needs for the RIA accelerators, in Proc. of 8th International Computational Accelerator Physics Conference, St. Petersburg, Russia, June 28-July 3, 2004.

[93] P. N. Ostroumov et al., Phys. Rev. ST/AB **5**, 060101 (2002).

[94] Advanced photon source upgrade, http://www.aps.anl.gov/News/Conferences/2006/APS_Upgrade/index.html.

[95] M. Borland, ERL upgrade options and possible performance, http://www.aps.anl.gov/News/Conferences/2006/APS_Upgrade/ERLUpgrades02.pdf, 2006.

[96] L. Merminga and *et al.*, ELIC design, http://casa.jlab.org/viewgraphs/merminga_EIC2004.pdf, 2004.

[97] W. Lu et al., Generating multi-GeV electron bunches using single stage laser wakefield acceleration in a 3D nonlinear regime, arxiv.org physics/0612227, 2006, Submmited to Physical Review STAB.

[98] C. Huang et al., Modeling TeV Class Plasmaa Fterburners, in *Proceedings of Particle Accelerator Conference*, pages 2666– 2668, 2005.

[99] S. Nagaitsev, T. Raubenheimer, and P. Tenenbaum, ILC computing and simulation needs for SciDAC proposal, Communication among ILC GDE advisory committee and SciDAC team, 2006.

[100] R. A. Rimmer et al., The JLab Ampere-class cryomodule, in Proc. of 12th SRF Workshop, Cornell University, July 10-15, 2005.

[101] G. Wu, M. Stirbet, H. Wang, and R. Rimmer, Multipacting analysis for JLAB ampere class cavities, in Proc. of 12th SRF Workshop, Cornell University, July 10-15, 2005.

[102] Cornell high energy synchrotron source, http://erl.chess.cornell.edu/.

[103] G. Bassi et al., Nucl. Instrum. Meth. **A557**, 189 (2006).

[104] David Bernholdt (PI), Technology for Advanced Scientific Component Software (TASCS), SciDAC2 CET, `http://www.scidac.gov/compsci/TASCS.html`, 2006.

[105] I. Zagorodnov, R. Schuhmann, and T. Weiland, Int. J. Numer. Model. **16**, 127 (2003).

A Committees and Collaboration list

In this appendix we list the membership of the committees responsible for the management and scientific direction of the project, together with the complete collaboration list.

Management Committee

John Cary	Tech-X Corporation
Lois Curfman McInnes	Argonne National Laboratory
Warren Mori	University of California, Los Angeles
Cho Ng	Stanford National Accelerator Laboratory
Esmond G. Ng	Lawrence Berkeley National Laboratory
Robert Ryne	Lawrence Berkeley National Laboratory
Panagiotis Spentzouris (Chair)	Fermi National Accelerator Laboratory

Scientific Advisory Committe

Ilan Ben-Zvi	Brookhaven National Laboratory
Michael Heath	University of Illinois at Urbana-Champaign
Chan Joshi	University of California at Los Angeles
Kwang-Je Kim	Argonne National Laboratory
Wim Leemans	Lawrence Berkeley National Laboratory
Lia Merminga	Thomas Jefferson National Laboratory
Sergei Nagaitsev	Fermi National Accelerator Laboratory
Steve Peggs (Chair)	Brookhaven National Laboratory
Tor Raubenheimer	Stanford Linear Accelerator Center
Vladimir Shiltsev	Fermi National Accelerator Laboratory
Valerie Taylor	Texas A&M University

COMPASS Collaboration

Argonne National Laboratory:	Michael Borland, Lois Curfman McInnes, Brahim Mustapha, Boyana Norris, Peter Ostroumov, Tim Tautges
Brookhaven National Laboratory:	Alexei Fedotov, Wolfram Fischer, Roman Samulyak
Fermi National Accelerator Laboratory:	James Amundson, Panagiotis Spentzouris, Eric Stern
Lawrence Berkeley National Laboratory:	E. Wes Bethel, William Fawley, Cameron Geddes, Xiaoye S. Li, Esmond G. Ng, Ji Qiang, Robert Ryne, Hongzhang Shan, Cristina E. Siegerist, Kurt Stockinger, Erich Strohmaier, Jean-Luc Vay
Los Alamos National Laboratory:	Peter Walstrom
Oak Ridge National Laboratory:	Richard Barrett
Stanford Linear Accelerator Center:	Volkan Akcelik, Andreas Kabel, Arno Candel, Lixin Ge, Kwok Ko, Lie-Quan Lee, Zenghai Li, Cho Ng, Ernesto Prudencio, Greg Schussman, Ravi Uplenchwar, Liling Xiao
State University of New York, Stonybrook:	Jim Glimm
Tech-X Corporation:	George Bell, David L. Bruhwiler, John Cary, Douglas Dechow, Andrey Sobol, Peter Stoltz
Thomas Jefferson National Accelerator Facility:	Rui Li, Haipeng Wang
University of California, Davis:	Kwan-Liu Ma
University of California, Los Angeles:	Viktor Decyk, Chengkun Huang, Warren Mori, Frank Tsung, Sven Reiche
University of Maryland:	Tom Antonsen
University of Southern California:	Tom Katsouleas

B Petascale

In electromagnetics, the main focus is on the use of petascale computing to model, design, and optimize accelerator components from the individual to the system level ultimately to optimize beam performance and physics output. An HEP exemplar of this is the Computational Grand Challenge of the International Linear Collider (ILC) Global Design Effort (GDE): "For a single three-cryomodule rf unit of the ILC Main Linac and by assuming realistic 3-D dimensions and misalignments, calculate multi-bunch beam dynamics effects, including wakefields and HOM excitations." Nuclear Physics will rely on energy recovery linacs (ERLs) for efficient electron coolers and for the more distant electron-ion collider (*c.f.* [96] and [4]). This will require the ability to model interactions with high-current beams, with cavity loading and re-entrance. A BES exemplar [chae06] comes from the proposed APS-2 upgrade based on an ERL, for which there is a need to compute dynamics of 100 fs bunches including the effects of both short- and long-range wakefields. In each case, there is a need to compute the cryogenics load induced by the HOMs, the consequences and/or mitigation of instabilities, and the optimization for larger stable currents and/or better energy recovery.

This team has developed complementary capabilities over the last several years. One approach has been to use higher order finite-element methods on unstructured meshes that conform to boundaries. Another approach has been oriented toward structured meshes using embedded boundaries. Our project defines a clear path of petascale scalability for both approaches, and we have identified applications where we can enable significant physics progress with petascale resources. For example, eigenmode computations of the complete three-cryomodule unit necessitates the solution of a large scale nonlinear complex eigenvalue problem with several hundred millions degrees of freedom (DOFs) using second order iso-parametric tetrahedral elements. For an exhaustive search of trapped modes in the rf unit, more than 3000 eigenmodes need to be computed. Given that the empirical time dependence is about $N^{1.5}$, where N is the number DOFs, about 100,000 wall-clock hours on a 1 TF computer with sufficient physical memory is required for one complete eigenmode analysis of the RF unit without including any algorithmic improvement. This would correspond to *100 wall-clock hours on a 1 petaflop computer*, assuming ideal scalability. Algorithm improvements with SAP researchers can further reduce the runtime significantly. With the structured mesh approach, the most brute force method of constant meshes has been shown to compute the electrodynamics of wake fields of single cavity modes with of the order of 1.3 billion degrees of freedom on a 0.5 TF cluster at the rate of roughly 1 hour per accelerating mode period. Petaflop computing would bring this down to seconds and ultimately enable full time-domain cryomodule modeling.

Under SciDAC2, the emphasis of beam dynamics modeling shifts toward multi-physics, multi-stage (first injection to extraction, then ultimately end to end) simulations. This could only become possible with the efficient utilization of petascale computing. Multi-physics, multi-stage simulations will enable the accelerator designers to quickly optimize design parameters, taking into account all relevant effects within the same model. They will also enable "nearline" modeling of operating accelerators, i.e simulation run time comparable to that of typical beam studies, thus providing feedback relevant to control room operations. For example, a simulation of one Fermilab Booster bunch including 3D space-charge and impedance effects, with realistic lattice and orbit control feedback, for 2,000 turns requires 20 hours on 512 processors on the NERSC SP3, using the Synergia framework. Since the computational effort is proportional to the number of turns $\times$ the number of bunches $\times$ the number of "kicks" in the simulation, a realistic simulation of the full cycle (20,000 turns) of all 84 bunches, relevant to a control room shift (8hrs) will require efficient utilization of 10,000 processors.

The efficient utilization of petascale computing will become possible with the development of better scaling, faster solvers, petascale efficient particle managers and load balancing algorithms, and further parallelization of the implemented physics models using pipelining. Of equal importance is the development of simulation frameworks that will allow optimal simultaneous interoperation of different physics modules on a petascle machine. To achieve the later we will build upon our SciDAC1 work on the Synergia and ML/I frameworks to provide a framework capable of utilizing the existing and future physics modules. For example, a realistic model of a machine such as the ILC DR from injection to extraction will require not only a multi-physics but also a multi-bunch simulation; such a model will require an order of magnitude higher computational effort than the Booster example discussed above.

The advanced accelerator community now sees clear paths for developing plasma based accelerator stages that could push the energy frontier of linear colliders. These paths rely on using physics that is nonlinear, therefore particle-in-cell (PIC) methods are required for high-fidelity modeling. Furthermore, the community has identified several grand challenge questions as described in section 1 that impact HEP and BES; answering these questions requires the use of PIC tools that run efficiently on petascale machines. For example, modeling a 1GeV LWFA stage using our explicit, SciDAC1, PIC codes on 256 processors takes 70,000 processor hours [97]. Modeling the L'Oasis 1GeV experiment will take 2,500,000 processor hours using an explicit PIC code because of the particulars of the plasma channel and laser spot size. In addition, it will take $\sim 5 \times 10^8$ cells, $\sim 10^9$ particles to model a 10 GeV stage using the full PIC codes OSIRIS or VORPAL. Such a simulation would take $\sim 2,500,000$ processor hours or ~ 1 day on a petaflop machine if these codes can scale and a 50 GeV LWFA stage might take a few weeks on a petaflop machine. When an external beam is to be injected or one requires some realtime feedback between simulation and experiment then a quasi-static PIC such as QuickPIC can be used. For example, it currently takes only 5,000 processor hours to model the L'Oasis experiment. Therefore, if QuickPIC can scale to a petaflop machine then the turn around time could be reduced to minutes and realtime feedback can be achieved. Modeling 50 GeV stages would require $\sim 100,000$ processor hours so if the code can scale then the turn around time for each run would be a few hours and parameter scans could be done.

Members of the SciDAC1 team have already simulated a 0.5 TeV afterburner stage using QuickPIC [98], but with larger transverse beam sizes than are envisaged. Such a simulation takes 5,000 node hours using the reduced PIC code QuickPIC (it would take over 5,000,000 processor hours using full PIC code). However, a 0.5TeV afterburner simulation using QuickPIC with proper resolution will require 10×10^9 cells and 6.4 million processor hours (assuming 3 Gflops/processor). On a petaflop machine this simulation could be done in $\sim$ a day.

With unprecedented realism and resolution achieved through petascale modeling, *Virtual Prototyping* (VP) of the design and optimization of accelerator components will revolutionize the current design process for accelerator builders. In a similar fashion, petascale computing will enable the development of a multi-physics beam dynamics simulation environment, a *Virtual Accelerator Model* (VAM), which will allow quantitative predictions at run times comparable to beam study run times. Through its interface with VP modules, the VAM will implement the most accurate representation of accelerator component characteristics and geometry. The VP on petascale computer platforms will substantially reduce the turnaround time for the *automatic* optimization procedure, realizing significant cost savings, while utilization of the VAM will provide the means for multi-parameter tunning and fast convergence of performance optimization.

C Target Applications of the Program Offices

Through our interactions with accelerator builders and designers, important problems of DOE/SC projects, the solution of which requires petascale computing, have been identified and prioritized. In addition to AM/CS support from SAP partners, the development of new capabilities in existing software and of novel application codes are required in order to address the complexity and diversity of the proposed problems. In the following we list the accelerator facilities that will benefit from the application activities of this project.

HEP applications

LHC: We will provide integrated beam dynamics modeling capabilities in support of LHC commissioning. Key required capabilities are fully nonlinear optics, beam-beam modeling, and electron-cloud modeling.

ILC: The ILC collaboration will be entering an intense phase of accelerator R&D in the coming years to produce a reference design report. The ILC GDE [99] has requested several specific tasks of our team. We will further develop our simulation capabilities to address these tasks so as to make a qualitative difference in the design with improved performance, increased reliability and lower cost.

(1) The ILC Computational Grand Challenge (CGC) - Determine wakefield effects for a single RF unit consisting of three crypmodules, each of which has eight baseline Tesla cavities, by taking into account of realistic cavity imperfections and misalignments. An important issue for such a design is trapped modes between cavities. Both cavity modes and trapped modes are affected by cavity imperfections (10s to 100s of μm) which are evident in measurements done at DESY on a cryomodule when compared with the results from an ideal cavity. These modes constitute the long-range wakefields that comprise the multi-bunch effects in the ILC bunch train. In a RF unit, we need to know for cooling purposes how much of the radiated HOM power is not absorbed by beamline dampers between the cryomodules. The problem is further complicated by cavity and cryomodule misalignments (100s of μm) which have to be taken into account. To respond to the challenge we will enhance our capabilities in the areas of linear and eigen solvers, shape determination and optimization, adaptive refinements, and visualization, in an integrated framework. A new finite-element based code TEM3P integrated thermal/electromagnetic/mechanical multi-physics solver will be developed to form a complete engineering prototyping toolset using one model within a unified framework. The wakefields will be interfaced to beam dynamics codes to assess multi-bunch effects on emittance preservation.

(2) Damping Ring (DR) - The successful operation of the damping rings is extremely important as they control the initial beam quality before it is injected into and then accelerated in the main Linac. Our SciDAC tools will allow the numerical model to be accurate to the engineering design level. There are three main objectives: a) to evaluate the short-range wakefields and provide a broadband impedance budget from all the vacuum chamber componets as input for beam dynamics instability studies, b) to identify all the trapped modes in the vacuum chamber components, c) to optimize performance and cost for the RF cavities and other vacuum chamber components.

In order to understand and optimize the performance of the machine, we will focus on multi-bunch, multi-physics effects simulations. The physics effects will include electron cloud, realistic geometry and impedances, IBS, higher-order magnetic optics, CSR, and space-charge. Ultimately, we will be able to perform simulations of the full cycle of the machine, from injection to extraction.

In addition to modeling electrom cloud effects, we will support the R&D effort to suppress them. We will implement surfacee impedance boundary condition for modeling grooved chamber walls used for suppressing electron cloud effects.

(3) Main Linac (ML) RF System - The Low-Loss (LL) and the ICHIRO cavities have been proposed as an alternative to the baseline TDR cavity to reduce the machine cost because of their higher gradients and lower cryogenic loss. We will focus on optimization of the end-groups to meet HOM damping requirements, the notch gap fields, and the coax pickup power flow. Multipacting is another limiting factor to high gradients in the ICHIRO cavity development and is potentially responsible for the long processing time of the TTF-III input coupler. A multi-mode waveguide boundary condition needs to be implemented in Omega3P to handle modes above beampipe cutoff for these cavities. Wakefields from cavity designs will be input into beam dynamic codes to study x-y coupling due to 3D effects such as cavity deformations and mode rotation.

(4) Beam Delivery System (BDS) - Control of wakefield effects from collimators, crab cavities and the interaction region (IR) in the BDS is critical to the successful operation of the machine. The issues related to crab cavities can be analysed using similar modeling techniques for the main Linac SRF cavities. Current FNAL design for the BDS crab cavity needs to be improved as the damping properties are not optimal while the HOM coupler tuning is too sensitive to small variations in the notch gap. Wakefield and heating effects in the IR will be analysed with higher resolution and increased accuracy. For the accurate determination of the short-range wakefield of collimators, a moving beam window using a non-dispersive algorithm in VORPAL and higher-order finite elements in T3P will be implemented and provide cross-checks of calculations. A Napoly-type wakefield integration technique needs to be implemented for collimator or intrusive structure in the vacuum chamber.

Another important issue for crab cavities is how cell-to-cell variation affects wakefields and coupling. To improve our ability to study such effects, we will add to the VORPAL code a new meshing capability that allows the user to specify a level of variation on a given baseline design.

We will also address HOM buildup and feedback on the beam for crab cavities, using a time-domain approach, which allows to study them simultaneously. Developing this time-domain approach has two benefits: (i) the approach is self-consistent, with the beam dynamics directly coupled to the electromagnetics, and (ii) we will be able to cross check the frequency domain approach for studying HOMs and beam feedback, since the two techniques are complementary.

In order to study emittance dilution and beam instabilities in the low emittance transport, the ML, and the interaction region, we will employ the same multi-physics, multi-bunch model described for the DR, with the addition of full beam-beam modeling, including quantum effects.

Tevatron RunII: We will focus on developing 3-D strong-strong beam-beam simulations of the motion of colliding bunches in the Tevatron, including impedance models. The simulations will include optics parameters reflecting the measured lattice and helical orbit. We will study possible coherent instabilities arising from the combined action of the beam-beam and impedance effects.

NuMI: We will perform e-cloud generation and e-cloud dynamics simulations for parameters relevant to the Main Injector (MI) proton plan (320 kW source with slip-stacking in MI). We will compare the predictions with MI measurements. The level of detail and sophistication of the simulations will increase during the course of the project, and ultimately we will perform multi-physics simulations. If the quality and quantity of the experimental data permits, the various comparisons will provide valuable insight on the range of applicability and limitations of the different model approximations. In addition, we will perform full (3D) beam dynamics simulations in recycler, including multi-particle effects, such as space-charge, for the recycler parameters envisioned by the SNuMI plan (super-beams for NuMI). In this scenario, the recycler is used as pre-injector for the MI.

Control Room Applications: We will continue to work with Fermilab Booster personnel to perform detailed Booster simulations, including space charge and impedance effects, and do comparisons with experimental data in order to build an accurate model of the machine. We will then use this detailed model to implement a prototype Control Room application. We will develop an interface to machine parameter and detector readouts (Ionization Profile Monitor, Beam Position Monitors, magnet currents, RF ramp) to implement an almost-real-time controls algorithm to help maximize intensity and efficiency. The application can them be ported to other machines of the FNAL proton source.

NP applications

RIA: We will optimize resonator design for the Rare Isotope Accelerator (RIA) Facility. RIA will be based on several types of normal-conducting Radio Frequency Quadrupole (RFQ) andsuperconducting TEM class resonators [91, 92, 93]. Stability of charged particles in RFQ, Hybrid RFQs and RFQ-DTLs resonators can be provided through dedicated geometry for accelerating gaps formed by drift tubes and focusing electrodes to exert the RF quadrupole focusing effect. In the injector section of the RIA post-accelerator the focusing by RF fields is extremely cost-effective and can save millions of dollars compared with any other alternative methods that can be applied for low charge-to-mass ratio ions with $q/A = 1/238$. much more efficient in low-velocity (beta < 0.7) region. For the optimization, integrated electromagnetic and mechanical analysis is required due to the complexity of the cavity geometry and mechanical properties.

CEBAF 12 GeV Upgrade: TJNAF faces challenging problems in SRF cavity design andmanufacturing for the CEBAF 12 GeV upgrade [100, 101]. These problems require high-accuracy 3D modeling of multi-cell cavities including end groups and couplers. There is great interest in integrated RF/thermal/stress capability to design high-power RF windows for low temperature structure cryo-cooling. Simulations of multipacting and dark currents will help identify problems during the high power processing of the superstructues, and those of space charge dominated effects will help evaluate the beam quality at the end of the injector. New waveguide boundary conditions need to be implemented in our frequency and time domain codes for proper termination at the rectangular HOM waveguide ports in modeling the cavities.

RHIC: BNL is pursing the development of a superconducting laser-photocathode RF gun (SRF-photoinjector) which is a novel device that holds the promise of generating a high-brightness, high-current continuous electron beam. The primary target is the RHIC electron cooler for NP, but its application can also be found in HEP (as a flat beam electron source), and in BES (for energy recovery linac (ERL) light sources). The design challenge of such a device is to produce relatively high charge (5 to 10 nC), high-brightness bunch with extremely good emittance. An accurate simulation requires a 3D code that includes wakefields, image charges, and retarded potentials and allows for a very large number of electrons in order to resolve short bunches (10 to 20 ps) starting off from the cathode non-relativistically. It is also necessary that the code is parallel so that one can realistically follow the evolution of the electron bunch through the accelerator until emittance compensation is established. To perform these simulations accurately and efficiently, the goal of this team is to develop a new parallel 3D finite element particle-in-cell (PIC) code Pic3P and improve the PIC and conformal boundary capabilities in VORPAL. The two codes will validate against each other for this new computational regime without possible known solution.

RHIC Upgrade, eRHIC, and ELIC : These devices will rely on electron cooling to reverse the heating caused by, e.g., beam-beam collisions. The electron cooling involves Energy Recover Linacs (ERLs) which extract the energy of the electrons after they have cooled the ion beams. Because the electrons are heated by this process, the energy recovery can be affected. Computations of the energy recovery from heated electron beams will be important for determining overall machine efficiency. Such computations will rely on self-consistent EM modeling, such as will be developed through the particle-in-cell methods for electromagnetic structures described above.

Electron colling simulations during the first three years will address the RHIC-II cooler, then turn to e-RHIC or ELIC parameters during the last two years. The parameter space to be explored is large, so we will use task-farming techniques to facilitate efficient parameter scans on emerging architectures with 10^5 processors.

We will also perform simulations of colliding beams, utilizing models of: a linac-ring and a circulator ring-ring collider, a crab cavity, a wire compensator, an electron lens, IBS, the space-charge effect of parallel co-moving beam. The new capabilities will be tested against analytical results and experimental data from RHIC. This will help to significantly improve the performance of the current RHIC operation, guide the design efforts of the planned RHIC upgrades, enhance the confidence level of the proposed eRHIC and ELIC design, and explore new parameter regimes.

BES applications

SNS: The SRF cavities in the SNS are experiencing problems with feed-through failures in the end groups. Large amount of power was extracted through the HOM coupler causing overheating. During this process, thermal radiation was observed near the end-groups, and particle emission from the cavity was suspected to be one cause. We will use realistic simulations, which include peak field and multipacting in the HOM coupler, to understand this phenomenon so that such failures could be prevented. In addition, we will model the effects of HOM damping to determine if the HOM couplers are really necessary for the cavities envisioned for the SNS power upgrade.

SNS researchers are already using IMPACT for modeling their linac and for modeling the transfer line from the linac to the accumulator ring. Under the present proposal we would continue to support our codes for SNS research, such as using Synergia and MaryLie/IMPACT for circular machine studies and for comparative studies with the code ORBIT.

APS upgrade: The APS upgrade [94] is considering many options, among which is an upgrade path to being an Energy Recovery Linac (ERL), similar to what was recently approved for development at Cornell [102]. However, there are [95] many uncertainties about feasibility, such as whether beam quality can be preserved, and whether the gun can sustain 100 mA. In addition, were it possible to reduce wakefield effects through 3D profile optimization, one could obtain greater currents and, therefore, x-ray fluxes as desired by users.

For creating the short electron bunches necessary for generating short x-ray pulses, researchers are considering crab cavities similar to those proposed for the ILC. Consequently, many of the electromagnetic research issues for the ILC, including HOM buildup and feedback onto the beam, also apply to the APS upgrade.

LCLS, future x-ray light sources: We will develop fast, high-fidelity, parallel computational capability for start-to-end design and analysis of 4th generation light sources. Such a model will have much in common with the computational tools developed for HEP and NP accelerator projects, including the need to model space-charge effects, wakefield effects, the ability to simulate tuning and commissioning algorithms, coupling to control rooms, etc. In addition, there are certain unique needs such as advanced parallel tools for modeling X-ray production. Light source development is an important area for BES, and a particularly timely one in light of the LCLS commissioning schedule. Also, it should be noted that BES has not previously been involved in a SciDAC accelerator modeling initiative. For these reasons, additional information on this portion of the COMPASS project is provided in the next Appendix.

D X-Ray Light Source Modeling for Ultrafast Science

The DOE Office of Basic Energy Sciences (BES) develops and operates nearly all of the nation's synchrotron light sources. These facilities are extraordinary tools for scientific research and discovery in fields such as materials science, chemistry, the biosciences, environmental science, and medicine. The coming fourth generation of light sources, based on X-ray free-electron lasers (XFEL's), offers new opportunities for studying chemical and biological reactions with unprecedented temporal resolution, but is also accompanied by significant technological challenges.

These challenges include producing and transporting high brightness, extremely low emittance electron beams; maintaining beam quality in the presence of space charge, wakefields, and coherent synchrotron radiation (CSR); and accurately predicting both coherent and incoherent undulator radiation. Understanding and predicting these phenomena is crucial for XFEL-based light sources. For example, uncontrolled growth of the microbunching instability can lead to irreversible degradation in electron beam quality, and ultimately reduce FEL performance; yet, presently this phenomenon is often modeled with highly simplified, 1D algorithms. Given the significant cost of the next-generation light sources, it is imperative that the DOE BES community have advanced, fast, high-fidelity parallel modeling capability for facility design as well as for the commissioning and operational stages of near-term facilities such as the LCLS.

The XFEL portion of this SciDAC-2 proposal differs from most of the others in that previously there was no SciDAC-1 support for code development for XFEL modeling applications. Of the codes involved in this proposal, only IMPACT was designed from the start as a parallel code. GINGER and GENESIS began as serial codes and evolved to take advantage of MPI-based "slice" parallelism. ELEGANT has long been used on clusters in a "concurrent computing" mode, but has just seen release of the first partially parallel version. However, few users (apart from the code authors) have exploited these parallel features, nor are all the codes mature in this area. This is a situation we plan to improve via the current proposal.

Because the code authors are predominantly accelerator/FEL scientists, we believe that we have a detailed and mature view of how improved codes can benefit light source design and operation. Thus, a large percentage of the initial proposed work is aimed toward relatively small scale (in terms of cost) extensions and interoperability that will pay large dividends in terms of light source optimization. An important example of this is the proposed work to give LCLS control room scientists a smoothly operational suite of start-to-end (i.e., cathode to FEL user station) computational tools that can model the entire system based on actual lattice settings, giving near real-time predictions of diagnostic output and FEL performance. Because of these operational, LCLS-oriented tasks, this proposal offers near-term benefits to BES. At the same time, it also includes other tasks that, through HEP/NP/BES collaboration under COMPASS, will result at the end of 5 years in a new generation of XFEL modeling tools in the hands of BES scientists.

Codes used for 4th Generation Light Source Modeling

We now briefly discuss the individual codes that form the backbone of this BES high brightness accelerator/XFEL proposal and then summarize the proposed work.

IMPACT: The IMPACT code suite, written in F90 and MPI, is widely used for modeling high intensity beams in rf proton and electron linacs and photoinjectors. It includes the 3D parallel PIC codes, IMPACT-T and IMPACT-Z. Together, these PIC codes allow large-scale, parallel simulation of a high brightness electron beam from cathode to undulator entrance. The code suite includes a number of features relevant to future light sources including self-consistent 3D space charge, ability to model high-aspect-ratio beams, wakes, and 1D CSR. The IMPACT suite has been used to model photoinjectors at BNL, FNAL, Cornell, and LCLS. It was used for the first 100M particle simulations of linac beam transport for a proposed soft X-ray light source (FERMI@Elettra).

ELEGANT: ELEGANT [10], developed at ANL, is a C-language, multi-platform code used for design, simulation, and optimization of FEL driver linacs, energy recovery linacs (ERLs), and storage rings. Researchers using ELEGANT to model LCLS were leaders in start-to-end simulations of linac-driven x-ray FELs, leading to discovery of a CSR-driven microbunching instability [11] and a subsequent redesign of the LCLS. ELEGANT was also used to verify theoretical predictions and cures for the longitudinal space charge-driven instability [12]. ELEGANT is used world-wide by essentially all FEL and ERL efforts that target the UV and x-ray wavelength regimes, including projects at Trieste, Frascati, DESY, Daresbury, BESSY, Cornell, JAERI, Pohang, etc. ELEGANT played important roles in ANL's LEUTL and BNL's VISA FEL projects, and was used to model the Stanford Subpicosecond Photon Source (SPPS).

GENESIS: GENESIS is a time-dependent, Fortran-based 3D FEL simulation code. It models single-pass free-electron lasers, both seeded and SASE FEL amplifier configurations; its flexible input has permitted its extension to model FEL oscillators and multistage harmonic cascades. GENESIS is in extensive worldwide use by nearly all major FEL projects, including the LCLS and, like GINGER, has been used to analyze many FEL experiments such as LEUTL and VISA. An MPI version of GENESIS was released in 2005 and currently runs on Linux clusters at UCLA and SLAC. GENESIS (and GINGER) import particle distributions from tracking codes such as ELEGANT to enable full start-to-end simulation capability.

GINGER: GINGER is a time-dependent, 2.5 dimensional FEL simulation code and currently runs in parallel mode on all major platforms at NERSC. Originally developed at LLNL in the mid-1980's to model high gain FEL amplifier experiments, GINGER was the first multidimensional FEL code to examine the spiky nature of SASE-based emission from the LCLS. Built on an F90/MPI source base, GINGER has been extended in numerous ways including the ability to model oscillator and harmonic cascade configurations and, most recently, to calculate harmonic radiation emission and utilize parallel HDF-5 I/O. Both GINGER and GENESIS have been benchmarked against each other for many situations, thus giving more confidence in their accuracy. Both codes treat undulator wakefields [18] and error fields, each of which can play important roles in XFEL's.

SPUR: SPUR calculates undulator spontaneous radiation emission in both the time and frequency domain. Based upon a Lienard-Wichert potential approach, the direction of observation is determined for each integration step, thus avoiding the far field approximation that would limit how close the observation point can be placed behind the undulator. In the MPI version, calculations for different observation points, defining the detector plane, are distributed over different CPU's. SPUR is currently used to estimate the undulator radiation from the LCLS XFEL, in part to help optimize the optical beamline for radiation damage threshold and background signal.

Highlights of Proposed Research

High Resolution, Multiscale Tracking Studies:
The design of a next-generation X-ray light source requires high fidelity modeling of a number of physical phenomena simultaneously, including nonlinear beam optics, space charge, wake fields, CSR, and others. Thanks to advances in parallel beam modeling under SciDAC1, we performed linac simulations using IMPACT-Z from 100 MeV to 1200 MeV for the FERMI@Elettra project using 100M simulation particles. Under the present proposal, we will link IMPACT-T (used to model the injector) and IMPACT-Z (used to model the remainder of the linac) to perform large

scale ($\geq$100M particle) simulations from cathode to undulator entrance for LCLS. Having fully parallelized the X-ray production codes, we will perform large-scale start-to-end simulations for LCLS. Along with these LCLS applications, we will also investigate new methods in the beam dynamics codes. We will explore a multi-scale approach involving use of a second, high resolution grid for a single slice of the bunch, and populate it with a large number of macroparticles. Such an approach, which will be valid when there is little or no particle movement between slices, will allow simulations with very low numerical shot noise within that slice for accurate studies of the microbunching instability.

Coherent Synchrotron Radiation (CSR) Modeling:
One focus of our proposed work is to enable fully 3D, efficient, high- fidelity simulations of CSR. This problem represents an immense algorithmic and computational challenge, and its solution will enable us to develop designs that control and mitigate the harmful effects of CSR. Currently, in fully self-consistent 3D simulations that employ particle- particle interactions, the calculations are so time consuming that only a small number of particles can be used. Consequently, most previous modeling efforts for such systems have involved some simplifications, with the net result that the impact of CSR for most short wavelength FEL systems cannot yet be reliably predicted in simulation. In this project we will utilize a method involving an EM mesh and Green function. However, we will use a different representation of the charge density compared with that used in CSRtrack [103], and we will use a parallel particle manager to reduce the computational bottleneck posed by interpolation. Our approach also replaces the brute force convolution with a mathematically equivalent $n \log n$ algorithm in the spatio-temporal gridsize. Using this approach we plan to perform parallel CSR simulations with more than million macroparticles.

X-Ray Production Modeling:
Many of the proposed tasks connected with x-ray calculation fall within the area of improving code-to-code data exchange interfaces (e.g. importing macroparticles from IMPACT to GENESIS/GINGER, ELEGANT to SPUR) and control room interfaces (tuning/optimization/experimental measurement prediction & analysis) discussed below. Other tasks are more oriented toward high performance computing: fuller parallelization of GENESIS, extending the GINGER grid and field solver to handle $r - \theta - z$ geometry, and then later exploring the feasibility of adding multigrid field solvers and better macroparticle loading schemes (e.g., Latin Hypercube Sampling) to both FEL codes. Another area is extending the parallelization of the SPUR code which calculates the spontaneous radiation from the undulator (which, importantly, for LCLS is 3 times more intense in total flux than is the coherent x-ray production). The output of GENESIS and SPUR will be conformed to the same format and combined if needed to model the X- ray optics beamline. Further SPUR improvements will include implementation of a fully 6D description of the electron beam distribution, the propagation of radiation wavefronts through optical elements (mirrors, gratings and aperture), and the determination of the effective photon distribution along the undulator for ray-tracing studies.

In the later years of this proposal, depending upon the then-current trends in XFEL R&D, we expect to investigate new, more self- consistent algorithms for XFEL radiation calculations. For very short (sub- femtosecond) radiation trains desired by experimentalists doing ultrafast science, the eikonal and discrete slippage approximations used in essentially all FEL codes start to become shaky. A more self-consistent field calculation may need to be developed and implemented, which would keep massive parallelism in mind from the ground up. We may also revisit the oscillator configuration coding in GINGER and GENESIS, in part to improve scalability and performance. Today, all sub-UV FEL's are single pass amplifiers due to the lack of highly reflective (and survivable) mirrors. However, by the end of this decade, this limitation may start to crumble and future soft-xray light source designers may need robust oscillator calculation capability. Here, the nature of FEL oscillators (e.g., multiple passes, new physics parameters such as desynchronism and hole outcoupling) implies that the computational intensity may actually be to or even larger that necessary to model a long but single pass amplifier such as the LCLS.

Tuning/Optimization/Commissioning Modeling:
Although current FEL designs are of course subjected to optimization, present procedures use approximate methods, often with simplified beam distributions, 2D tracking codes, or approximate analytical expressions. Once fast, reliable, parallel tools are available for start-to-end modeling, we can consider start-to-end optimization of such a system. To accomplish this, we would first identify critical parameters of the design (e.g., photoinjector laser temporal shape, beam compressor settings) and then identify key performance parameters, such as x-ray pulse length and energy and create a suitable penalty function. The optimization itself could use one of several techniques such as the genetic technique, multi-objective technique [20], or another technique [21]. Following this, a more conventional optimization technique could be applied to refine the design found using the approximate model. Obviously this is an area of broad programmatic applicability that would be carried out in close collaboration with other COMPASS participants from HEP and NP.

Coupling to Control Rooms:
This is an area that also has broad applicability across the COMPASS project. Among the capabilities provided by such a system would be:

- An easy-to-use graphical user interface (GUI). This is essential in a control room environment where rapid results are needed to make efficient use of expensive beam time.

- Use of fast parallel multi-physics modeling of beam transport and x-ray production, with the user able to seamlessly choose faster, less accurate models or slower, but more accurate models.

- Modeling of the present state of the accelerator with output that permits direct comparison with control room diagnostic measurements (e.g., beam position and profile monitors, output radiaton spectra, etc.)

- Modeling of experiments, i.e., parameter scans. This ability would be available at all phases of an experiment, allowing accelerator physicists both to plan better experiments and then later to better analyse and understand the results.

- Fitting beam distribution models to diagnostic measurements. This would include not only fitting the linear optics via response matrix measurements, but also fitting the 6D beam distribution to match the results of transverse and longitudinal phase space and radiation measurements.

E Embedded Scientific Application Partnerships

As introduced in Section 3.1, all applied mathematics and computer science work in this project supports the broader computational physics goals of the team, with direct applications to the offices of HEP, NP, and BES. We organize this ASCR-supported work according to embedded Scientific Application Partnership (SAP) activities, which are collaborations among physicists on the accelerator proposal team and members of the following SciDAC2 Centers for Enabling Technology (CET) and Institutes:

- Towards Optimal Petascale Simulations (TOPS) [80]
- Center for Interoperable Technologies for Advanced Petascale Simulations (ITAPS) [81]
- Applied Partial Differential Equations Center (APDEC) [83]
- Combinatorial Scientific Computing and Petascale Simulations (CSCAPES) [82]
- Technology for Advanced Scientific Component Software (TASCS) [104]
- Performance Evaluation Research Institute (PERI) [85]
- SciDAC Institute for Ultra-Scale Visualization (IUSV) [86]
- Visualization and Analytics Center for Enabling Technology (VACET) [87]

We request embedded SAP funding for researchers who will interact with the accelerator physics team and members of the various CETs and Institutes. Thus, our SAP portfolio will facilitate bidirectional collaboration, with computational accelerator physicists motivating and validating the tools and techniques under development by the CETs and Institutes, and the math and computer science teams enhancing the productivity of the computational physicists. Letters of collaboration are provided in Appendix I. The following SAP topics are presented in the remainder of this appendix, with emphasis on the prior accomplishments, motivation, approach, and impact of each activity:

E.1 Development of Novel Scalable Parallel Eigensolvers for Large Electromagnetic Cavity Simulations (with TOPS)

E.2 Shape Determination and Optimization for Superconducting Cavities (with TOPS and ITAPS)

E.3 Domain Specific Scalable Linear Solvers for Finite-Element Large-Scale Accelerator Cavity Simulations (with TOPS)

E.4 Development of a Super-Fast Poisson Solver for Accelerator Physics Applications (with TOPS)

E.5 Parallel Adaptive Refinement for High-Fidelity Finite-Element Electromagnetic System Simulations (with ITAPS and CSCAPES)

E.6 Embedded Boundary Methods for Particle-in-Cell Simulations of Electromagnetic structures (with ITAPS)

E.7 The UPIC Framework (with APDEC, TOPS, and PERI)

E.8 Performance Analysis and Optimization for Accelerator Simulation Codes at Petascale Supercomputers (with PERI)

E.9 Interactive and Remote Visualization for Massive Complex Field-Particle Data on Unstructured Grids (with IUSV)

E.10 High-Performance Parallel Data Management and Analysis Tools for Beam Dynamics Modeling (with VACET and SDM)

E.11 Computational Quality-of-Service Enabled Beam Dynamics Modeling (with TASCS, TOPS, and PERI)

E.1 Development of Novel Scalable Parallel Eigensolvers for Large Electromagnetic Cavity Simulations

Investigators: Esmond Ng (LBNL)
Physics Collaborators: Cho Ng and Kwok Ko (SLAC)
Math/CS Collaborators: Esmond Ng (LBNL/TOPS)

Prior Accomplishments. Eigenmode analysis is one of the most vital yet powerful tools for electromagnetic cavity design, analysis, and optimization. Under SciDAC1, the filter algorithm (a combination of inexact shift-invert Lanczos and Jacobi-Davidson method) [sun2003] was developed as a result of a Stanford Ph.D. thesis work embedded at SLAC. The algorithm enabled us to carry out successfully many electromagnetic cavity designs and modeling such as the NLC cell design with 0.1% accuracy [li2001] and the heating study of PEP-II IR [li2004b]. With the use of state-of-art sparse direct solvers, we implemented an efficient parallel implementation of the shift-invert Lanczos method, which allowed us to successfully model and analyze the 55-cell detuned structure (DS) and damped detuned structure (DDS) of the NLC [li2004a, li2004b]. After superconducting technology had been chosen for the ILC, we quickly developed and implemented several methods, including the Self-Consistent Iteration, the Second-Order Arnoldi [bai2005, lee2005a], and the Nonlinear Jacobi-Davidson methods [vos2004b, lia2006], for the complex nonlinear eigensystems in the eigemode analysis of the ILC cavities [lee2005a, lee2005b, lee2005c, lee2005d].

Proposed Work

Motivation. Mode analysis of waveguide-load SRF cavities requires the solution of the vector wave equation with a boundary condition at a waveguide port that is a nonlinear function of the cavity resonance frequency. This boundary condition leads to a nonlinear term in an otherwise linear eigenvalue problem in the finite element discretization procedure [lee2005a, lee2005b]. The real part of an eigenvalue is a cavity resonance frequency, and the imaginary part represents the damping due to waveguide ports. The eigenvector corresponds to the electromagnetic field distribution in the cavity. A mode is trapped inside the cavity when the imaginary part of the eigenvalue is very small. Trapped modes in extended superconducting RF structures (ILC cryomodules, superstructures, and RF units of the main Linac) may cause beam instability or excessive cryogenics loss. An exhaustive search for the trapped modes requires the solution of nonlinear eigenvalue problems with hundred million degrees of freedom (DOFs). With the existing algorithms, it either requires prohibitively large amount of memory per node or excessively long execution time, which makes it infeasible to obtain scientific results.

Approach. Under the support of SciDAC2, we will improve the existing eigensolvers and develop novel scalable eigensolvers on petascale supercomputers to tackle the above challenging and complex problems. Some approaches include, but are not limited to, developing the algebraic multi-level substructuring (AMLS) technique [yan2005, els2005] for linear and nonlinear eignenvalue problems, further exploring the iterative projection method for the nonlinear eigenvalue problem, developing nonlinear eigensolvers specific to multi-mode waveguide-loaded cavities, and exploring alternative efficient eigensolvers on the petascale platforms. This effort will be performed in close collaboration with the Towards Optimal Petascale Simulations (TOPS) Center. Our focus will be on the integration and optimization of the eigenvalue solvers for large electromagnetic cavity simulation.

Impact. The development of novel algorithms or significant enhancement to existing eigensolvers for petascale computers is indispensable to solving computationally challenging problems in cavity design, optimization and analysis. The advances in eigensolver research for large-scale simulations will provide an unparallel capability for accelerator scientists to answer compelling questions such as the beam stability by realistic modeling of extended accelerator structures at the system level.

Milestones

The activities will be performed in close collaboration with TOPS.

Year 1

- Incorporate TOPS algebraic multi-level substructuring (AMLS) solver for linear eigenvalue problems into Omega3P.
- Enhance the performance and scalability of the nonlinear Arnoldi and nonlinear Jacobi-Davidson algorithms to solve large-scale nonlinear eigenvalue problems in ILC multi-cavity cryomodule simulations.

Year 2

- Incorporate TOPS algebraic multi-level substructuring (AMLS) solver for nonlinear eigenvalue problems into Omega3P.
- Evaluate the performance of parallel AMLS for linear and nonlinear eigenvalue problems for petascale computing platforms.

Year 3

- Develop parallel nonlinear eigensolver specific to multi-mode waveguide-loaded cavities.
- Incorporate TOPS scalable subspace projection methods for nonlinear eigenvalue problems into Omega3P for petascale computing platforms.

Year 4

- Harden TOPS scalable linear and nonlinear eigensolvers used in electromagnetic simulations on petascale computing platforms.
- Investigate alternative eigensolvers for extremely large-scale linear and nonlinear eigenvalue problems for electromagnetic simulations on petascale systems.

Year 5

- Continue work on alternative parallel linear and nonlinear eigensolvers for extremely large-scale electromagnetic simulations.

References

[bai2005] Z. Bai, et. al., SOAR, A Second-order Arnoldi Method for the Solution of the Qaudratic Eigenvalue Problem, SIAM Journal on Matrix Analysis and Applications, Vol. 26, No. 3, pp 640-659, 2005.

[els2005] Kolja Elssel, Heinrich Voss, Automated Multilevel Substructuring for Nonlinear Eigenproblems, Report 86, Arbeitsbereich Mathematik, TU Hamburg-Harburg 2005. Proc. Internat. Conf. on Civil, Structural, and Engineering Computing, Rome, Italy, 2005.

[lee2005a] Lie-Quan Lee, Zhaojun Bai, Weiguo Gao, Lixin Ge, Parry Husbands, Marc Kowalski, Xiaoye Li, Zenghai Li, Cho-Kuen Ng, Chao Yang, Esmond Ng, and Kwok Ko, Modeling RF Cavities with External Coupling, SIAM Conference on Computational Science & Engineering, Orlando, Florida, February 12-15, 2005.

[lee2005b] Lie-Quan Lee, Zhaojun Bai, Weiguo Gao, Lixin Ge, Zenghai Li, Cho-Kuen Ng, Chao Yang, Esmond Ng, and Kwok Ko, Nonlinear Eigenvalue problem in Acceleartor Cavity Design, SIAM Annual Meeting, New Orleans, Louisiana, July 11-15, 2005.

[lee2005c] Lie-Quan Lee, Lixin Ge, Zenghai Li, Cho-Kuen Ng, Greg Schussman, and Kwok Ko, Achievements in ISICs/SAPP Collaborations for Electromagnetic Modeling of Accelerators, in Proceedings of SciDAC 2005 Conference, San Francisco, California, 2005.

[lee2005d] L. Lee, et. al., Modeling RF Cavities with External Coupling SIAM Conference on Computational Science and Engineering, Orlando, Florida.

[li2001] Zenghai Li, Nathan Folwell, Gene Golub, Adam Guetz, Brian McCandless, Cho-Kuen Ng, Yong Sun, Michael Wolf, Ron Yu and Kwok Ko, Parallel Computations of Wakefields in Linear Colliders and Storage Rings, Proceedings of the Particle Accelerator Conference, Chicago, Illinois, June 18-22, 2001.

[li2004a] Zenghai Li, Nathan T. Folwell, Lixin Ge, Adam Guetz, Valentin Ivanov, Marc Kowalski, Lie-Quan Lee, Cho-Kuen Ng, Greg Schussman, Ravindra Uplenchwar, Michael Wolf, and Kwok Ko, X-band Linear Collider R&D in Accelerating Structures Through Advanced Computing, in Proceedings of the 9th European Particle Accelerator Conference, Lucerne, Switzerland, July 5-9, 2004.

[li2004b] Zenghai Li, Nathan Folwell, Lixin Ge, Adam Guetz, Valentin Ivanov, Marc Kowalski, Lie-Quan Lee, Cho-Kuen Ng, Greg Schussman, Ravi Uplenchwar, Michael Wolf, Liling Xiao, and Kwok Ko, High Performance Computing in Accelerator Structure Design and Analysis, in Proceedings of the 8th International Computational Accelerator Physics Conference, St. Petersburg, Russia, June 29-July 2, 2004.

[lia2006] Ben-Shan Liao, et. al., An Iterative Projection Method For Solving Large-Scale Nonlinear Eigenproblems with Application to Next-Generation Accelerator Design, Tenth Copper Mountain Conference on Iterative Methods, Colorado, April 2-7, 2006.

[sun2003] Yong Sun, The Filter Algorithm for Solving Large-Scale Eigenproblems from Accelerator Simulations, Ph.D. thesis, Stanford University.

[yan2005] C. Yang, W. G. Gao, Z. Bai, X. Li, L. Lee, P. Husbands and E. G. Ng, An Algebraic Sub-structuring Algorithm for Large-scale Eigenvalue Calculation, SIAM J. Sci. Comp., vol. 27, no. 3, pages 873-892, 2005.

[vos2004a] H. Voss, An Arnoldi method for nonlinear eigenvalue problems, BIT Numerical Mathematics 44, pp387, 2004.

[vos2004b] H. Voss, A Jacobi-Davidson Method for Nonlinear Eigenproblems, Proceedings of ICCS04, Krakow, Poland, pp34-41, 2004.

E.2 Shape Determination and Optimization for Superconducting Cavities

Investigators: Volkan Akcelik (SLAC), Tim Tautges (ANL/ITAPS)
Physics Collaborators: Cho Ng and Kwok Ko (SLAC)
Math/CS Collaborators: Omar Ghattas (Univ. of Texas, Austin/TOPS)

Prior Accomplishments. In SciDAC1, a collaboration involving TOPS, TSTT, and SLAC initiated an effort to develop a parallel, automatic shape optimization and a shape determination capability for accelerating cavity design and for cavity shape determination using measured data, respectively. Both shape optimization and determination were mathematically posed as PDE constrained optimization problems [akc2005,akc2006]. In the mathematical optimization procedure, the design variables for shape optimization were the CAD parameters, and those for shape determination the unknown cavity deformations. The PDE constraint was the Maxwell equation in the frequency domain. The optimization algorithm was based on a gradient-based approach, and the design sensitivities were computed using an adjoint based method [pir1983]. In the prototyped SLAC code, the adjoint problem was discretized with the same finite element method used in SLAC application codes and solved with existing solvers. Mesh smoothing and design velocities required for gradient computations were performed analytically for a specified geometry. The code was successfully applied to a simplified ILC cavity and provided the optimized cavity shape with the correct fundamental frequency of 1.3 GHz [lee2005c]. In parallel, TSTT partners prototyped a CAD-based procedure for more complex models which required numerical evaluations of design velocities in CAD models and numerical mesh smoothing in subsequent changes of geometry [tau2005].

Proposed Work

Motivation. The main part of shape optimization and shape determination will be carried out in SciDAC2 by the same research team working together before under SciDAC1. By bringing together expertise in optimization (TOPS), in advanced linear and eigen solvers (TOPS), in adaptive mesh control and mesh smoothing (ITAPS), and in cavity design (SLAC), an advanced procedure involving the aforementioned components will be implemented through the application of state-of-the-art algorithms in the area of large-scale PDE constrained optimization. This capability, when completed, will drastically reduce the time and effort in optimizing cavities that are required to satisfy multiple designed objective functions and constraints. While shape optimization improves the cavity shape design, the inverse problem of shape determination allows a cavity shape with imperfections to be computed from measured data. This is useful for modeling superconducting RF cavities used in accelerators such as CEBAF, SNS and at Cornell, and those planned for the ILC and RIA. Due to the tolerance errors during the cavity fabrication process and mechanical tuning of the cavity for recovering the correct cavity operating mode frequency, the cavity final shape will be different from its designed shape. The determination of the real dimensions of the deformed cavity is crucial in evaluating the effects of imperfections on beam stability in accelerators.

Approach. We plan to implement the CAD-based prototypes into a fully parallel optimization code. The CAD-based mesh smoothing tools and the numerical calculations of design velocities developed by ITAPS will be incorporated into the gradient-based optimization formulation. The code will run on parallel computers (NERSC Seaborg and Bassi, NCCS Phoenix and DOE petascale computers when available) and will be applied to the most challenging computational problems such as the ILC complex system of superconducting cavities. A prototype shape determination code [akc2006b], which was developed during SciDAC1, will be modified for the shape determination of the realistic ILC cavity. In the previous prototype code, the damping of unwanted cavity modes has been ignored, and the forward optimization is a real and linear eigenvalue problem. The damping effects from couplers in the actual cavity will transform the forward optimization to a complex nonlinear eigenvalue problem, the solution of which will be facilitated by two accompanied SAP proposals on nonlinear eigensolvers and parallel linear solvers. The optimization code will be used to determine the true dimensions of ILC cavities from measured data by solving an inverse problem. The true cavity shape will then be used for physics studies.

Impact. The successful development of an optimization code in finite element electromagnetic system simulation of superconducting cavities under SciDAC will undoubtedly be of its first kind at petascale computing. It requires a concerted multidisciplinary effort in the areas of applied mathematics, computer science, and physics application from two SciDAC CETs and accelerator laboratories. As a matter of fact, a Gordon-Bell prize winner formerly supported by TOPS under SciDAC1 will carry out the optimization work at SLAC. This demonstrates that computational and physical scientists from various disciplines can be integrated into a coherent team for the advancement of computational science. While the shape optimization and determination work is targeted to a challenging problem of national scientific significance clearly related to DOE missions through computational science, the tools developed will equally benefit accelerator designers in the Offices of BES, HEP and NP in improving the performance of their machines.

Milestones for SLAC

The activities will be performed in close collaboration with TOPS and ITAPS CETs.

Year 1

- Incorporate ITAPS geometry/mesh deformation and design velocity capabilities into Omega3P.
- Integrate TOPS optimization software in shape determination and optimization for unloaded cavities.

Year 2

- Integrate TOPS optimization software in shape determination and optimization for open, waveguide-loaded cavities.
- Evaluate/improve techniques for computing design velocities for superconducting RF cavities.

Year 3

- Investigate the effects of control variables on the computed geometry of a cavity using the shape determination software.
- Assemble and evaluate a comprehensive parallel tool for shape determination and optimization of superconducting cavities using solver and meshing components from TOPS and ITAPS on petascale computers.

Year 4

- Validate the parallel shape determination and optimization tool for superconducting cavities.
- Improve the parallel performance of the different components of shape determination and optimization code on the petascale compute platform.

Year 5

- Harden the parallel shape determination and optimization tool for superconducting cavities.
- Further improve the optimization and determination capabilities.

Milestones for ANL

Year 1

- Improve ITAPS mesh movement capabilities to include high order tetrahedral elements.
- Investigate the parallel strategies for mesh movement and design velocity computations.

Year 2

- Validate ITAPS meshing tool DDRIV for mesh movement and design velocities in electromagnetic simulations.
- Improve the runtime performance of the meshing tool through detailed analysis.

Year 3

- Improve the usability and quality of mesh movement and design velocity computations.
- Optimize the memory usage in the meshing tool.

Year 4

- Explore the need of adaptive mesh refinement in the shape determination and optimization tool.

Year 5

- Further enhance the parallel meshing tool for the petascale computer platforms.

References

[akc2006] V. Akcelik, G. Biros, O. Ghattas, J. Hill, D. Keyes, and B. van Bloemn Waanders, ”Parallel algorithms for PDE-constrained optimization”, in Frontiers of Parallel Computing, M. Heroux, P. Raghaven, and H. Simon, eds, SIAM, 200, to appear.

[akc2005] Volkan Akcelik, George Biros, Omar Ghattas, David Keyes, Kwok Ko, Lie-Quan Lee, and Esmond Ng, Adjoint Methods for Electromagnetic Shape Optimization of the Low-Loss Cavity for the International Linear Collider, in Proceedings of SciDAC 2005 Conference, San Francisco, California, 2005.

[akc2006b] V. Akcelik et al., ”Parallel Adjoint-Based Methods for Electromagnetic Shape Optimization”, SIAM Ccnference on Parallel Processing for Scientific Computing, San Francisco, California, February 22-24, 2006.

[pir1983] O. Pironneau, ”Optimal Shape Design for Elliptic Systems”, Springer-Verlag, 1983.

[lee2005c] Lie-Quan Lee, Lixin Ge, Zenghai Li, Cho-Kuen Ng, Greg Schussman, and Kwok Ko, Achievements in ISICs/SAPP Collaborations for Electromagnetic Modeling of Accelerators, in Proceedings of SciDAC 2005 Conference, San Francisco, California, 2005.

[tau2005] T.J. Tautges, P. Knupp, J.A. Kraftcheck, H.J. Kim, Interoperable Geometry and Mesh Components for SCIDAC Applications, in Proceedings of SciDAC 2005, Scientific Discovery Through Advanced Computing, volume 16 of J. Physics: Conference Series, Institute of Physics Publishing, 2005.

E.3 Domain Specific Scalable Linear Solvers for Finite-Element Large-Scale Accelerator Cavity Simulations

Investigators: Lie-Quan Lee (SLAC)
Physics Collaborators: Cho Ng and Kwok Ko (SLAC)
Math/CS Collaborators: Esmond Ng (LBNL/TOPS)

Prior Accomplishments. Fast linear solvers are essential in electromagnetic system simulations of accelerator cavities and components using SLAC parallel finite-element code suite Omega3P, S3P, T3P, and Pic3P. In SciDAC1, we made three major accomplishments in the area of parallel linear solvers. First, by using a new joint vector-scalar potential scheme [de1996] that we developed for the representation of finite element basis functions, the effective condition number of the matrix has been reduced and the convergence of the iterative linear solvers has been improved drastically. For typical simulations, the number of iterations using a preconditioned conjugate gradient method in the new scheme is reduced by a factor of 10, and subsequently the computation time by a factor of six [ko2004]. The second accomplishment is the development of a generic linear solver framework which enables the use of either direct solvers (e.g. WSMP, SuperLU, and MUMPS) or iterative methods (Iterative Template Library) in solving very large sparse linear systems [ko2004]. The third accomplishment is the development of a parallel hierarchical preconditioner [lee2004] that enables the simulation of electromagnetic systems with a hundred million degrees of freedom (DOF). Using this hierarchical preconditioner, a real linear system with 4 billion non-zeros and 93 million unknowns from 3D electromagnetic simulations was solved in less than 8 minutes with 1024 CPUs on the NERSC Seaborg supercomputer [lee2004]. This unprecedented simulation capability led us to the first-ever direct wakefield end-to-end simulation of the 55-cell detuned structure and damped-detuned structure of the Next Linear Collider [ko2004, li2004a]. When there is enough physical memory, the use of direct solvers provides 50-100 times faster solution time than the use of iterative solvers in calculating multiple tightly-clustered interior eigenvalues [ko2004].

Proposed Work

Motivation. In simulating the ILC cryomodule and the entire RF unit of the ILC main Linac using Omega3P, S3P, and T3P (ILC Computational Grand Challenge), we will need to repeatedly solve extremely large sparse linear systems. In the case of nonlinear eigenvalue computations with existing solvers such as Second Order Arnoldi and Nonlinear Jacobi-Davidson methods, the shifted linear systems are complex and symmetric. Applying sparse direct linear solvers to such large-scale linear systems will require a prohibitively large amount of per-node memory to store the matrix factors while using iterative methods often leads to excessively long execution time for solving the linear systems with multiple right hand sides. This poses a big computational challenge for the exhaustive search for all the unwanted trapped modes in the ILC multi-cavity cryomodules and RF units.

Approach. In SciDAC2, we propose to use domain specific methods to enhance the existing linear solvers and to develop new scalable linear solvers on petascale platforms. We will study the effectiveness of finite-element based multilevel preconditioner for solving the shifted complex linear systems in the solution of nonlinear eigenvalue problems. We will develop memory-reducing techniques, such as using single-precision version of the sparse direct solvers with double-precision iterative refinement in place of double-precision version of the sparse direct solvers [lee2007]. We will search for better ordering schemes for preserving sparsity in sparse matrix factorization and look into the approach of using out-of-core storage for the factors in the sparse direct solvers. In the context of the shape determination and optimization for ILC cavities, we need to solve a structured indefinite linear system - the Karush-Kuhn-Tucker (KKT) linear system,

$$\begin{pmatrix} \mathbf{K} - \lambda\mathbf{M} & \mathbf{Mv} \\ (\mathbf{Mv})^T & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{t} \\ \xi \end{pmatrix} = \begin{pmatrix} \mathbf{b} \\ \mathbf{0} \end{pmatrix} \tag{1}$$

where $(\lambda, \mathbf{v})$ is the eigenpair of the matrix pencil $(\mathbf{K}\ , \mathbf{M})$, and the (1, 1)-block numerically singular with the null space $\mathbf{v}$. We will apply sparse direct solvers and Krylov subspace iterative methods to the KKT linear systems associated with cavity shape determination and optimization. We will develop effective domain specific preconditioners for solving the KKT linear system, and algorithms that will take advantage of the known null space of the sub-matrix. The proposed work will be carried out in close collaboration with TOPS researchers on their activities in linear solvers.

Impact. The successful advancement of the domain specific scalable linear solvers along with petascale computers will enable physicists to design, model, and optimize realistic accelerator structures to a level of complexity, speed, and accuracy orders of magnitude beyond what is currently possible. It will help accelerator physicists to address challenging issues such as simulating the entire RF unit of ILC main Linac in HEP and multi-cavity simulations and modeling proposed in NP and BES.

Milestones

The activities will be performed in close collaboration with TOPS.

Year 1
- Develop effective domain-specific finite-element based multilevel preconditioners for simulating multi-cavity systems.
- Deploy TOPS sparse direct solvers for solving KKT linear systems associated with shape determination and optimization problems.

Year 2
- Deploy TOPS Krylov subspace iterative methods and effective preconditioners for solving KKT linear systems associated with shape determination and optimization problems.
- Enhance the capabilities of sparse direct solvers for large-scale electromagnetic application codes by means of memory-reducing techniques.

Year 3
- Collaborate with TOPS to resolve memory scalability issues in sparse direct solvers for large-scale electromagnetic simulations.
- Enhance the scalability of linear solvers on petascale platforms for both frequency-domain and time-domain large-scale simulations for the RF unit of ILC main Linac.

Year 4
- Integrate new parallel sparsity-preserving ordering schemes for sparse matrix factorization from TOPS into electromagnetic application codes.
- Continue to improve the performance and scalability (both time and memory usage) of the linear solvers.

Year 5

- Explore algorithms that will take advantage of the known null space of the (1, 1)-block in the KKT system.
- Evaluate the efficiency of TOPS parallel out-of-core sparse direct solvers in electromagnetic simulations.

References

[**de1996**] Roman Dyczij-Edlinger and Oszkar Biro, A Joint Vector and Scalar Potential Formulation for Driven High Frequency Problems Using Hybrid Edge and Nodal Finite Elements, IEEE Transaction on Microwave Theory and Techniques, Vol. 44, No. 1, PP 15-23, 1996.

[**ko2004**] Kwok Ko et al., Electromagnetic Systems Simulation and Prototyping through Computation, SciDAC Report, 2004.

[**lee2004**] Lie-Quan Lee, Lixin Ge, Marc Kowalski, Zenghai Li, Cho-Kuen Ng, Greg Schussman, Michael Wolf, and Kwok Ko, Solving Large Sparse Linear Systems in End-to-end Accelerator Structure Simulations, in Proceedings of the 18th International Parallel and Distributed Processing Symposium, Santa Fe, New Mexico, April 26-30, 2004.

[**lee2007**] Lie-Quan Lee, Volkan Akcelik, Lixin Ge, Ernesto Prudencio, and Kwok Ko, On Reducing Memory Usage of the Simulations Using Sparse Direct Solvers, to be presented on 2007 SIAM Conference on Computational Science and Engineering, Costa Mesa, California, 2007.

[**li2004a**] Zenghai Li, Nathan T. Folwell, Lixin Ge, Adam Guetz, Valentin Ivanov, Marc Kowalski, Lie-Quan Lee, Cho-Kuen Ng, Greg Schussman, Ravindra Uplenchwar, Michael Wolf, and Kwok Ko, X-band Linear Collider R&D in Accelerating Structures through Advanced Computing, in Proceedings of the 9th European Particle Accelerator Conference, Lucerne, Switzerland, July 5-9, 2004.